Lecture Notes in Computer Science 4556

Commenced Publication in 1973
Founding and Former Series Editors:
Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Editorial Board

David Hutchison
Lancaster University, UK

Takeo Kanade
Carnegie Mellon University, Pittsburgh, PA, USA

Josef Kittler
University of Surrey, Guildford, UK

Jon M. Kleinberg
Cornell University, Ithaca, NY, USA

Friedemann Mattern
ETH Zurich, Switzerland

John C. Mitchell
Stanford University, CA, USA

Moni Naor
Weizmann Institute of Science, Rehovot, Israel

Oscar Nierstrasz
University of Bern, Switzerland

C. Pandu Rangan
Indian Institute of Technology, Madras, India

Bernhard Steffen
University of Dortmund, Germany

Madhu Sudan
Massachusetts Institute of Technology, MA, USA

Demetri Terzopoulos
University of California, Los Angeles, CA, USA

Doug Tygar
University of California, Berkeley, CA, USA

Moshe Y. Vardi
Rice University, Houston, TX, USA

Gerhard Weikum
Max-Planck Institute of Computer Science, Saarbruecken, Germany

Constantine Stephanidis (Ed.)

Universal Access in Human-Computer Interaction

Applications and Services

4th International Conference on Universal Access
in Human-Computer Interaction, UAHCI 2007
Held as Part of HCI International 2007
Beijing, China, July 22-27, 2007
Proceedings, Part III

Volume Editor

Constantine Stephanidis
Foundation for Research and Technology - Hellas (FORTH)
Institute of Computer Science (ICS)
70013 Heraklion, Crete, Greece
E-mail: cs@ics.forth.gr

Library of Congress Control Number: 2007930218

CR Subject Classification (1998): H.5.2, H.5.3, H.3-5, C.2, K.4.2, I.3, D.2, F.3

LNCS Sublibrary: SL 2 – Programming and Software Engineering

ISSN 0302-9743
ISBN-10 3-540-73282-9 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-73282-2 Springer Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media

springer.com

© Springer-Verlag Berlin Heidelberg 2007
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 12082612 06/3180 5 4 3 2 1 0

Foreword

The 12th International Conference on Human-Computer Interaction, HCI International 2007, was held in Beijing, P.R. China, 22–27 July 2007, jointly with the Symposium on Human Interface (Japan) 2007, the 7th International Conference on Engineering Psychology and Cognitive Ergonomics, the 4th International Conference on Universal Access in Human-Computer Interaction, the 2nd International Conference on Virtual Reality, the 2nd International Conference on Usability and Internationalization, the 2nd International Conference on Online Communities and Social Computing, the 3rd International Conference on Augmented Cognition, and the 1st International Conference on Digital Human Modeling.

A total of 3403 individuals from academia, research institutes, industry and governmental agencies from 76 countries submitted contributions, and 1681 papers, judged to be of high scientific quality, were included in the program. These papers address the latest research and development efforts and highlight the human aspects of design and use of computing systems. The papers accepted for presentation thoroughly cover the entire field of Human-Computer Interaction, addressing major advances in knowledge and effective use of computers in a variety of application areas.

This volume, edited by Constantine Stephanidis, contains papers in the thematic area of Universal Access in Human Computer Interaction, addressing the following major topics:

- Web and Media Accessibility and Usability
- Universal Access to Information and Communication
- Learning and Entertainment
- Universal Access to eServices

The remaining volumes of the HCI International 2007 proceedings are:

- Volume 1, LNCS 4550, Interaction Design and Usability, edited by Julie A. Jacko
- Volume 2, LNCS 4551, Interaction Platforms and Techniques, edited by Julie A. Jacko
- Volume 3, LNCS 4552, HCI Intelligent Multimodal Interaction Environments, edited by Julie A. Jacko
- Volume 4, LNCS 4553, HCI Applications and Services, edited by Julie A. Jacko
- Volume 5, LNCS 4554, Coping with Diversity in Universal Access, edited by Constantine Stephanidis
- Volume 6, LNCS 4555, Universal Access to Ambient Interaction, edited by Constantine Stephanidis
- Volume 8, LNCS 4557, Methods, Techniques and Tools in Information Design, edited by Michael J. Smith and Gavriel Salvendy
- Volume 9, LNCS 4558, Interacting in Information Environments, edited by Michael J. Smith and Gavriel Salvendy
- Volume 10, LNCS 4559, HCI and Culture, edited by Nuray Aykin

- Volume 11, LNCS 4560, Global and Local User Interfaces, edited by Nuray Aykin
- Volume 12, LNCS 4561, Digital Human Modeling, edited by Vincent G. Duffy
- Volume 13, LNAI 4562, Engineering Psychology and Cognitive Ergonomics, edited by Don Harris
- Volume 14, LNCS 4563, Virtual Reality, edited by Randall Shumaker
- Volume 15, LNCS 4564, Online Communities and Social Computing, edited by Douglas Schuler
- Volume 16, LNAI 4565, Foundations of Augmented Cognition 3rd Edition, edited by Dylan D. Schmorrow and Leah M. Reeves
- Volume 17, LNCS 4566, Ergonomics and Health Aspects of Work with Computers, edited by Marvin J. Dainoff

I would like to thank the Program Chairs and the members of the Program Boards of all Thematic Areas, listed below, for their contribution to the highest scientific quality and the overall success of the HCI International 2007 Conference.

Ergonomics and Health Aspects of Work with Computers

Program Chair: Marvin J. Dainoff

Arne Aaras, Norway
Pascale Carayon, USA
Barbara G.F. Cohen, USA
Wolfgang Friesdorf, Germany
Martin Helander, Singapore
Ben-Tzion Karsh, USA
Waldemar Karwowski, USA
Peter Kern, Germany
Danuta Koradecka, Poland
Kari Lindstrom, Finland
Holger Luczak, Germany
Aura C. Matias, Philippines
Kyung (Ken) Park, Korea
Michelle Robertson, USA
Steven L. Sauter, USA
Dominique L. Scapin, France
Michael J. Smith, USA
Naomi Swanson, USA
Peter Vink, The Netherlands
John Wilson, UK

Human Interface and the Management of Information

Program Chair: Michael J. Smith

Lajos Balint, Hungary
Gunilla Bradley, Sweden
Hans-Jörg Bullinger, Germany
Alan H.S. Chan, Hong Kong
Klaus-Peter Fähnrich, Germany
Michitaka Hirose, Japan
Yoshinori Horie, Japan
Richard Koubek, USA
Yasufumi Kume, Japan
Mark Lehto, USA
Robert Proctor, USA
Youngho Rhee, Korea
Anxo Cereijo Roibás, UK
Francois Sainfort, USA
Katsunori Shimohara, Japan
Tsutomu Tabe, Japan
Alvaro Taveira, USA
Kim-Phuong L. Vu, USA
Tomio Watanabe, Japan
Sakae Yamamoto, Japan

Jiye Mao, P.R. China
Fiona Nah, USA
Shogo Nishida, Japan
Leszek Pacholski, Poland
Hidekazu Yoshikawa, Japan
Li Zheng, P.R. China
Bernhard Zimolong, Germany

Human-Computer Interaction

Program Chair: Julie A. Jacko

Sebastiano Bagnara, Italy
Jianming Dong, USA
John Eklund, Australia
Xiaowen Fang, USA
Sheue-Ling Hwang, Taiwan
Yong Gu Ji, Korea
Steven J. Landry, USA
Jonathan Lazar, USA
V. Kathlene Leonard, USA
Chang S. Nam, USA
Anthony F. Norcio, USA
Celestine A. Ntuen, USA
P.L. Patrick Rau, P.R. China
Andrew Sears, USA
Holly Vitense, USA
Wenli Zhu, P.R. China

Engineering Psychology and Cognitive Ergonomics

Program Chair: Don Harris

Kenneth R. Boff, USA
Guy Boy, France
Pietro Carlo Cacciabue, Italy
Judy Edworthy, UK
Erik Hollnagel, Sweden
Kenji Itoh, Japan
Peter G.A.M. Jorna, The Netherlands
Kenneth R. Laughery, USA
Nicolas Marmaras, Greece
David Morrison, Australia
Sundaram Narayanan, USA
Eduardo Salas, USA
Dirk Schaefer, France
Axel Schulte, Germany
Neville A. Stanton, UK
Andrew Thatcher, South Africa

Universal Access in Human-Computer Interaction

Program Chair: Constantine Stephanidis

Julio Abascal, Spain
Ray Adams, UK
Elizabeth Andre, Germany
Margherita Antona, Greece
Chieko Asakawa, Japan
Christian Bühler, Germany
Noelle Carbonell, France
Jerzy Charytonowicz, Poland
Pier Luigi Emiliani, Italy
Michael Fairhurst, UK
Zhengjie Liu, P.R. China
Klaus Miesenberger, Austria
John Mylopoulos, Canada
Michael Pieper, Germany
Angel Puerta, USA
Anthony Savidis, Greece
Andrew Sears, USA
Ben Shneiderman, USA
Christian Stary, Austria
Hirotada Ueda, Japan

Gerhard Fischer, USA
Jon Gunderson, USA
Andreas Holzinger, Austria
Arthur Karshmer, USA
Simeon Keates, USA
George Kouroupetroglou, Greece
Jonathan Lazar, USA
Seongil Lee, Korea
Jean Vanderdonckt, Belgium
Gregg Vanderheiden, USA
Gerhard Weber, Germany
Harald Weber, Germany
Toshiki Yamaoka, Japan
Mary Zajicek, UK
Panayiotis Zaphiris, UK

Virtual Reality

Program Chair: Randall Shumaker

Terry Allard, USA
Pat Banerjee, USA
Robert S. Kennedy, USA
Heidi Kroemker, Germany
Ben Lawson, USA
Ming Lin, USA
Bowen Loftin, USA
Holger Luczak, Germany
Annie Luciani, France
Gordon Mair, UK
Ulrich Neumann, USA
Albert "Skip" Rizzo, USA
Lawrence Rosenblum, USA
Dylan Schmorrow, USA
Kay Stanney, USA
Susumu Tachi, Japan
John Wilson, UK
Wei Zhang, P.R. China
Michael Zyda, USA

Usability and Internationalization

Program Chair: Nuray Aykin

Genevieve Bell, USA
Alan Chan, Hong Kong
Apala Lahiri Chavan, India
Jori Clarke, USA
Pierre-Henri Dejean, France
Susan Dray, USA
Paul Fu, USA
Emilie Gould, Canada
Sung H. Han, South Korea
Veikko Ikonen, Finland
Richard Ishida, UK
Esin Kiris, USA
Tobias Komischke, Germany
Masaaki Kurosu, Japan
James R. Lewis, USA
Rungtai Lin, Taiwan
Aaron Marcus, USA
Allen E. Milewski, USA
Patrick O'Sullivan, Ireland
Girish V. Prabhu, India
Kerstin Röse, Germany
Eunice Ratna Sari, Indonesia
Supriya Singh, Australia
Serengul Smith, UK
Denise Spacinsky, USA
Christian Sturm, Mexico
Adi B. Tedjasaputra, Singapore
Myung Hwan Yun, South Korea
Chen Zhao, P.R. China

Online Communities and Social Computing

Program Chair: Douglas Schuler

Chadia Abras, USA
Lecia Barker, USA
Amy Bruckman, USA
Peter van den Besselaar, The Netherlands
Peter Day, UK
Fiorella De Cindio, Italy
John Fung, P.R. China
Michael Gurstein, USA
Tom Horan, USA
Piet Kommers, The Netherlands
Jonathan Lazar, USA
Stefanie Lindstaedt, Austria
Diane Maloney-Krichmar, USA
Isaac Mao, P.R. China
Hideyuki Nakanishi, Japan
A. Ant Ozok, USA
Jennifer Preece, USA
Partha Pratim Sarker, Bangladesh
Gilson Schwartz, Brazil
Sergei Stafeev, Russia
F.F. Tusubira, Uganda
Cheng-Yen Wang, Taiwan

Augmented Cognition

Program Chair: Dylan D. Schmorrow

Kenneth Boff, USA
Joseph Cohn, USA
Blair Dickson, UK
Henry Girolamo, USA
Gerald Edelman, USA
Eric Horvitz, USA
Wilhelm Kincses, Germany
Amy Kruse, USA
Lee Kollmorgen, USA
Dennis McBride, USA
Jeffrey Morrison, USA
Denise Nicholson, USA
Dennis Proffitt, USA
Harry Shum, P.R. China
Kay Stanney, USA
Roy Stripling, USA
Michael Swetnam, USA
Robert Taylor, UK
John Wagner, USA

Digital Human Modeling

Program Chair: Vincent G. Duffy

Norm Badler, USA
Heiner Bubb, Germany
Don Chaffin, USA
Kathryn Cormican, Ireland
Andris Freivalds, USA
Ravindra Goonetilleke, Hong Kong
Anand Gramopadhye, USA
Sung H. Han, South Korea
Pheng Ann Heng, Hong Kong
Dewen Jin, P.R. China
Kang Li, USA
Zhizhong Li, P.R. China
Lizhuang Ma, P.R. China
Timo Maatta, Finland
J. Mark Porter, UK
Jim Potvin, Canada
Jean-Pierre Verriest, France
Zhaoqi Wang, P.R. China
Xiugan Yuan, P.R. China
Shao-Xiang Zhang, P.R. China
Xudong Zhang, USA

In addition to the members of the Program Boards above, I also wish to thank the following volunteer external reviewers: Kelly Hale, David Kobus, Amy Kruse, Cali Fidopiastis and Karl Van Orden from the USA, Mark Neerincx and Marc Grootjen from the Netherlands, Wilhelm Kincses from Germany, Ganesh Bhutkar and Mathura Prasad from India, Frederick Li from the UK, and Dimitris Grammenos, Angeliki Kastrinaki, Iosif Klironomos, Alexandros Mourouzis, and Stavroula Ntoa from Greece.

This conference could not have been possible without the continuous support and advise of the Conference Scientific Advisor, Gavriel Salvendy, as well as the dedicated work and outstanding efforts of the Communications Chair and Editor of HCI International News, Abbas Moallem, and of the members of the Organizational Board from P.R. China, Patrick Rau (Chair), Bo Chen, Xiaolan Fu, Zhibin Jiang, Congdong Li, Zhenjie Liu, Mowei Shen, Yuanchun Shi, Hui Su, Linyang Sun, Ming Po Tham, Ben Tsiang, Jian Wang, Guangyou Xu, Winnie Wanli Yang, Shuping Yi, Kan Zhang, and Wei Zho.

I would also like to thank for their contribution towards the organization of the HCI International 2007 Conference the members of the Human Computer Interaction Laboratory of ICS-FORTH, and in particular Margherita Antona, Maria Pitsoulaki, George Paparoulis, Maria Bouhli, Stavroula Ntoa and George Margetis.

April 2007

Constantine Stephanidis
General Chair, HCI International 2007

HCI International 2009

The 13th International Conference on Human-Computer Interaction, HCI International 2009, will be held jointly with the affiliated Conferences in San Diego, California, USA, in the Town and Country Resort & Convention Center, 19-24 July 2009. It will cover a broad spectrum of themes related to Human Computer Interaction, including theoretical issues, methods, tools, processes and case studies in HCI design, as well as novel interaction techniques, interfaces and applications. The proceedings will be published by Springer. For more information, please visit the Conference website: http://www.hcii2009.org/

General Chair
Professor Constantine Stephanidis
ICS-FORTH and University of Crete
Heraklion, Crete, Greece
Email: program@hcii2009.org

Table of Contents

Part I: Web and Media Accessibility and Usability

Accessibility Guidelines Management Framework 3
Myriam Arrue, Markel Vigo, Amaia Aizpurua, and Julio Abascal

Accessibility Evaluation for Multimedia Content 11
Chieko Asakawa, Takashi Itoh, Hironobu Takagi, and Hisashi Miyashita

Tenuta: Strategies for Providing Guidance on Usability and Accessibility 20
Nigel Bevan, Helen Petrie, and Nigel Claridge

Remote Navigability Testing Using Data Gathering Agents 28
Martin Gonzalez-Rodriguez, Fidel Diez-Diaz, and Agueda Vidau

Enabling Accessibility and Enhancing Web Experience: Ordering Search Results According to User Needs 34
Alice Good and Jenny Jerrams-Smith

An Evaluation of Accessibility of Hierarchical Data Structures in Data Broadcasting -Using Tactile Interface for Visually-Impaired People- 45
Takuya Handa, Tadahiro Sakai, Kinji Matsumura, Yasuaki Kanatsugu, Nobuyuki Hiruma, and Takayuki Ito

Web Accessibility Compliance of Government Web Sites in Korea 55
Joonho Hyun, Doojin Choi, and Sukil Kim

Usability Evaluation of Korean e-Government Portal 64
Seongil Lee and Joo Eun Cho

Accessibility Evaluation of Korean e-Government 73
Seongil Lee, Byung Geun Kim, and Jong Gon Kim

The Potential of Web Accessibility in China: A Hypothesis on Its Impact on the Global Web Interface 79
Eleanor Lisney, Christina Li, and Songtao Liu

Restoring Semantics to BML Content for Data Broadcasting Accessibility 88
Kinji Matsumura, Yasuaki Kanatsugu, Takuya Handa, and Tadahiro Sakai

Making Multimedia Internet Content Accessible and Usable 98
Hisashi Miyashita, Hironobu Takagi, Daisuke Sato, and Chieko Asakawa

An Evaluation of Mobile 3D-Based Interaction with Complex Multimedia Environments 108
Ali A. Nazari Shirehjini and Jafar Movahedi

Detection of Layout-Purpose TABLE Tags Based on Machine Learning 116
Hidehiko Okada and Taiki Miura

The Usability of Accessibility Evaluation Tools 124
Helen Petrie, Neil King, Carlos Velasco, Henrike Gappa, and Gaby Nordbrock

Accessibility Works: Enhancing Web Accessibility in Firefox 133
John T. Richards, Vicki L. Hanson, Jonathan P. Brezin, Calvin B. Swart, Susan Crayne, and Mark R. Laff

Improving Accessibility for Existing Websites Spanning Multiple Domains 142
Takashi Sakairi, Takuya Ohko, and Hidemasa Muta

Large Scale Web Accessibility Evaluation - A European Perspective 150
Mikael Snaprud and Agata Sawicka

The BenToWeb XHTML 1.0 Test Suite for the Web Content Accessibility Guidelines 2.0 - Last Call Working Draft 160
Christophe Strobbe, Jan Engelen, Johannes Koch, Carlos Velasco, Evangelos Vlachogiannis, and Daniela Ortner

DHTML Accessibility Checking Based on Static JavaScript Analysis.... 167
Takaaki Tateishi, Hisashi Miyashita, Tabuchi Naoshi, Shin Saito, and Kouichi Ono

A Unified Web Evaluation Methodology Using WCAG 177
Eric Velleman, Christophe Strobbe, Johannes Koch, Carlos A. Velasco, and Mikael Snaprud

Improving Accessibility Through the Visual Structure of Web Contents 185
Masahiro Watanabe, Daisuke Asai, and Yoko Asano

Editing a Test Suite for Accessibility of Interactive Web Sites 193
Gerhard Weber and Kurt Weimann

Part II: Universal Access to Information and Communication

Discrimination and Perception of the Acoustic Rendition of Texts by Blind People 205
Vassilis Argyropoulos, Konstantinos Papadopoulos, Georgios Kouroupetroglou, Gerasimos Xydas, and Philippos Katsoulis

Using Speech Recognition and Intelligent Search Tools to Enhance Information Accessibility 214
Keith Bain, Jason Hines, Pawan Lingras, and Yumei Qin

Critical Success Factors for Automatic Speech Recognition in the Classroom 224
Steve Bennett, Jill Hewitt, Barry Mellor, and Caroline Lyon

Developing and Evaluating a TriAccess Reading System 234
Ming-Chung Chen, Chien-Chuan Cko, Lih-Yuan Chen, and Chun-Han Chiang

CAT Motor: An Innovative System to Detect the Behavior of Human Computer Interaction for People with Upper Limb Impairment 242
Chien-Chuan Cko, Ming-Chung Chen, Ting-Fang Wu, Su-Ying Chen, and Chih-Ching Yeh

The BIC Model: A Blueprint for the Communicator 251
J.P. de Ruiter and N.J. Enfield

Conveying Browsing Context Through Audio on Digital Talking Books 259
Carlos Duarte and Luís Carriço

An RFID-Based Point-and-Listen Interface Providing Library Access for the Visually Impaired 269
Daniel Fallman, Kent Lindbergh, Oskar Fjellstrom, Lars Johansson, Fredrik Nilbrink, and Linda Bogren

Performance Evaluation of Voice Interaction as a Universal Web Interface 279
Yu-Ming Fei, Chiuhsiang Joe Lin, Min-Ting Chen, and Chih-Cheng Chiang

Fisheye Views of Java Source Code: An Updated LOD Algorithm 289
J. Louise Finlayson, Chris Mellish, and Judith Masthoff

A Bayesian Network Approach to Semantic Labelling of Text Formatting in XML Corpora of Documents 299
Florendia Fourli-Kartsouni, Kostas Slavakis, Georgios Kouroupetroglou, and Sergios Theodoridis

Winsight: Towards Completely Automatic Backtranslation of Nemeth Code 309
D. Gopal, Q. Wang, G. Gupta, S. Chitnis, H. Guo, and A. Karshmer

Authoring Tools for Structuring Text-Based Activities 319
Maria Grigoriadou and Grammatiki Tsaganou

VoxBoox: A System for Automatic Generation of Interactive Talking Books 329
Aanchal Jain and Gopal Gupta

An On-Screen Keyboard for Users with Poor Pointer Control 339
Rick Kjeldsen

Multimodal Technology for Municipal Inspections: An Evaluation Framework 349
Irina Kondratova, Jeff Rankin, and Ashley Goggin

The Effects of Spatially Enriched Browsing Shortcuts on Web Browsing of Blind Users 358
Christos Kouroupetroglou, Michail Salampasis, and Athanasios Manitsaris

Chatting with an AAC(Augmentative/Alternative Communication Device) 368
Scott H. Kuebler and Ronald C. Kuebler

A Flexible On-Screen Keyboard: Dynamically Adapting for Individuals' Needs 371
Yun-Lung Lin, Ming-Chung Chen, Ya-Ping Wu, Yao-Ming Yeh, and Hwa-Pey Wang

Evaluation of RSVP and Display Types on Decoding Performance of Information Extraction Tasks 380
Ya-Li Lin and Chun-Min Ho

A Large Scale Study of English-Chinese Online Dictionary Search Behavior 389
Yong Liu and Jianmiao Fan

Design of Computer Integrated Safety and Health Management System 399
Hunszu Liu

Sensor-Driven Adaptation of Web Document Presentation 406
Spyros Nathanail, Vassileios Tsetsos, and Stathes Hadjiefthymiades

Effectiveness of Communication Process Support for Senior Citizens with Information Machines 416
Yosoko Nishizawa, Yoshinori Wakizaka, Takayuki Yamamoto, and Harumi Hasebe

Technology Advances and Standardization Toward Accessible Business Graphics 426
Hironobu Takagi and Tatsuya Ishihara

Auditory Accessibility of Metadata in Books: A Design for All Approach 436
Dimitrios Tsonos, Gerasimos Xydas, and Georgios Kouroupetroglou

Enhancing the Usability of Real-Time Speech Recognition Captioning Through Personalised Displays and Real-Time Multiple Speaker Editing and Annotation 446
Mike Wald and Keith Bain

A Study on the Straight-Line Drawing Tasks for the Non-sighted People 453
Chih-Fu Wu, Cherng-Yee Leung, Yung-Hsiang Tu, Kuang-Chieh Hsu, , and Huei-Tang Lin

Performance of Different Pointing Devices on Children with Cerebral Palsy 462
Ting-Fang Wu and Ming-Chung Chen

Part III: Learning and Entertainment

User Modeling for Intelligent Interfaces in e-Learning 473
Ray Adams

Ubiquitous Access to Learning Material in Engineering 481
Andreas Auinger and Christian Stary

ICT Educational Tools and Visually Impaired Students: Different Answers to Different Accessibility Needs 491
Stefania Bocconi, Silvia Dini, Lucia Ferlino, Cristina Martinoli, and Michela Ott

Modeling Content and Expression of Learning Objects in Multimodal Learning Management Systems 501
Giorgio Brajnik

Entertainment and Ambient: A New OLDES' View 511
M. Busuoli, T. Gallelli, M. Haluzík, V. Fabián, D. Novák, and O. Štěpánková

Virtual and Mixed Reality Interfaces for e-Training: Examples of Applications in Light Aircraft Maintenance 520
Johannes Christian, Horst Krieger, Andreas Holzinger, and Reinhold Behringer

A Holistic Approach to the Evaluation of E-Learning Systems 530
M.F. Costabile, T. Roselli, R. Lanzilotti, C. Ardito, and V. Rossano

How Can HCI Factors Improve Accessibility of m-Learning for Persons with Special Needs? 539
Matjaž Debevc, Mateja Verlič, Primož Kosec, and Zoran Stjepanovič

Accessible and Adaptive e-Learning Materials: Considerations for Design and Development 549
Matjaž Debevc, Zoran Stjepanovič, Petra Povalej, Mateja Verlič, and Peter Kokol

Web 2.0 Technology: Future Interfaces for Technology Enhanced Learning? 559
Martin Ebner, Andreas Holzinger, and Hermann Maurer

Computer Interface for Learning and Using Virtual Instrument 569
Yongkai Fan, Tianze sun, Jun Lin, Xiaolong Fu, and Yangyi Sui

Agents in m-Learning Systems Based on Intelligent Tutoring 578
Vlado Glavinic, Marko Rosic, and Marija Zelic

Digital Game-Based Learning: An Agent Approach 588
Christian Anthony L. Go and Won-Hyung Lee

A Haptic Based Interface to Ease Visually Impaired Pupils' Inclusion in Geometry Lessons 598
Cédric Gouy-Pailler, Sophie Zijp-Rouzier, Sylvie Vidal, and Denis Chêne

Unified Design of Universally Accessible Games 607
Dimitris Grammenos, Anthony Savidis, and Constantine Stephanidis

Universal Design Within the Context of e-Learning 617
Andrina Granić and Maja Ćukušić

Designing Intelligent Interfaces for e-Learning Systems: The Role of User Individual Characteristics 627
Andrina Granić and Jelena Nakić

Providing Universal Accessibility Using Connecting Ontologies: A Holistic Approach 637
Shuaib Karim, Khalid Latif, and A. Min Tjoa

Immersive Digital Games: The Interfaces for Next-Generation E-Learning? 647
Michael D. Kickmeier-Rust, Neil Peirce, Owen Conlan, Daniel Schwarz, Dominique Verpoorten, and Dietrich Albert

Methodology and Application in a Technologically Enhanced Environment: The Case of the Writing Center at the Hellenic American Union and the Hellenic American University 657
Vassiliki Kourbani, Yiannis Petropoulos, and Dimitris Tolias

The Role of Web-Based Learning Environments in Fostering Collaboration 667
Jasna Kuljis and Lorna Lines

The Virtual Paddle: An Universal Interaction for Accessible Video Games 677
Guillaume Lepicard, Frédéric Vella, Nadine Vigouroux, Benoît Rigolleau, David Chautard, and Elian Pucheu

Learning by e-Learning: Breaking Down Barriers and Creating Opportunities for the Visually-Impaired 687
Barbara Leporini and Marina Buzzi

Designing Games to Address 'Mute English' Among Children in China 697
Jason Li, Neema Moraveji, Jiarong Ding, Patrick O'Kelley, and Suze Woolf

Identification of Threat of Isolation as a Result of Work with a Computer in Free Time 707
Teresa Musioł and Katarzyna Ujma-Wąsowicz

Reaching Beyond the Invisible Barriers: Serving a Community of Users with Multiple Needs 716
Dianna L. Newman and Gary Clure

Some Aspects of the Development of Low-Cost Augmented Reality Learning Environments as Examples for Future Interfaces in Technology Enhanced Learning 728
Alexander Nischelwitzer, Franz-Josef Lenz, Gig Searle, and Andreas Holzinger

Students' Attidudes Towards Novel Interfaces in E-Learning 738
Margit Pohl, Ilona Herbst, Franz Reichl, and Sylvia Wiltner

The Impact of m-Learning in School Contexts: An "Inclusive" Perspective 748
Francesca Pozzi

Instructional Technology and Faculty Development: How iWRITE Challenges Course Design and Teaching Methods 756
Margaret Procter

Technology-Mediated Provision of Models: Observational Learning for First Year Students 766
J. Barbara Rose

English for Academic Purposes in Israel: Perceptions of E-Learning from the Perspectives of Learners and Teachers 776
Bella Rubin and Helen Sarid

AudioStoryTeller: Enforcing Blind Children Reading Skills 786
Jaime Sánchez and Iván Galáz

Basic Skills Training to Disabled and Adult Learners Through an Accessible e-Learning Platform 796
Olga C. Santos, Jesús G. Boticario, Alicia Fernández del Viso, Santiago Pérez de la Cámara, Carlos Rebate Sánchez, and Emmanuelle Gutiérrez y Restrepo

Interface of Online Mini-Go-Game with Pen Input Guide for the Blind 806
Michio Shimizu, Masahiko Sugimoto, and Kazunori Itoh

Enhancing Universal Access – EEG Based Learnability Assessment 813
Christian Stickel, Josef Fink, and Andreas Holzinger

Learning Through Exploration, Autonomy, Collaboration, and Simulation: The 'all-in-one' Virtual School of the Hellas Alive!© Online, Language-Learning Platform 823
Dimitris Tolias and George Exadaktylos

Co-learn: Collaborative Learning Engine a Vision for Pervasive Collaboration in E-Learning 833
Vaibhav Tyagi and Sheikh Iqbal Ahamed

Effects of Group Composition in Collaborative Learning of EFL Writing 843
Siew-Rong Wu

Part IV: Universal Access to eServices

E- Health System for Coagulation Function Management by Elderly People 853
Noemi Biterman, Eyal Lerner, and Haim Bitterman

Feed the Dragon Wisely: Designing for Childhood Awareness as a Means of Lifelong Obesity Prevention 860
Shunying Blevis, Jeffrey Bardzell, and Nancy Wroblewski

2D Numeric-Based Voice-Driven Browser for the Visual Impairments . . . 869
Chi Nung Chu, Yu Ting Huang, and Yao-Ming Yeh

Case Study of Applying Wireless Technologies into Healthcare Industry in China and UK 874
Guixia Kang, Li Zhang, Shanghong Li, Ping Zhang, and Said Boussakta

Attitude Toward Wireless Applications in Healthcare Industry in China 883
Guixia Kang, Yue Ouyang, Da Liu, Huaqing Wang, and Ping Zhang

Users and Trust: The New Threats, the New Possibilities 893
Kristiina Karvonen

Ontology-Based User Preference Modeling for Enhancing Interoperability in Personalized Services 903
Ju-Yeon Kim, Jong-Woo Kim, and Chang-Soo Kim

A Non-visual Approach to Improving Collaboration Between Blind and Sighted Internet Users 913
Ravi Kuber, Wai Yu, and Graham McAllister

Idea Movement of Aging Citizens: Lessons-Learnt from Innovation Workshops 923
Jaana Leikas

Accessing e-Government Services: Design Requirements for the Older User 932
Lorna Lines, Oluchi Ikechi, and Kate S. Hone

An Accessible Platform for Conference Administration and Management 941
George Margetis, Stavroula Ntoa, Maria Bouhli, and Constantine Stephanidis

Service Science: Applications of Observation to Real World Business 951
Haruhito Matsunami

VoiceBlog for Blind and Weak-Eyed People 961
Yoshie Sagata, Masahiro Watanabe, and Yoko Asano

Inflection of In-Situ Ergonomics in Service Engineering 970
Atsushi Shinya

Distributive Medical Management System 979
Barbara Sorensen, Judith Riess, and Eric Allely

Basic Consideration for Various Interfaces from Viewpoint of Service . . . 990
A. Toshiki Yamaoka and Mayuko Yoshida

Specialized Design of Web Search Engine for the Blind People 997
Yi-Fan Yang and Sheue-Ling Hwang

The Augmented Patient Chart: Seamless Integration of Physical and Digital Artifacts for Hospital Work 1006
Myrna S. Zamarripa, Victor M. Gonzalez, and Jesus Favela

Author Index 1017

Accessibility Guidelines Management Framework

Myriam Arrue, Markel Vigo, Amaia Aizpurua, and Julio Abascal

Laboratory of Human-Computer Interaction for Special Needs, Informatika Fakultatea, University of the Basque Country, Manuel Lardizabal 1, 20018 Donostia, Spain
myriam@si.ehu.es, markel@si.ehu.es, aaizpurua009@ikasle.ehu.es, julio@si.ehu.es

Abstract. Due to the extensive amount of information regarding web accessibility developed in recent years, developers find difficulties for creating accessible web applications. Accessibility knowledge management tools can facilitate this task. However, existing tools do not provide enough support to all the activities of the development process. In this sense, developers have to use diverse tools focused on different stages and they are also required to search for accessibility information from different sources. The aim of this paper is to present a framework for managing accessibility guidelines. It serves as a central repository where developers can search for accessibility guidelines, define new guidelines, and share them with other developers and evaluating them automatically. Therefore, this framework facilitates developers' tasks as it gives support to several activities of different stages throughout the development process.

1 Introduction

In recent years, an extensive amount of information about web accessibility in terms of best practices, techniques and sets of guidelines has been developed. However, some web developers are not able to produce accessible web content and most of the existing web sites present several accessibility barriers. In fact, web accessibility aware community complains about the amount of heterogeneous guidelines and the difficulty when handling them. Due to the complexity of the accessible web applications development task, web developers usually have to deal with diverse problems [1]:

- Search for the sets of guidelines which are significant for the current development.
- Select the most adequate sets of guidelines.
- Verify the coherence of the selected sets of guidelines.
- Analyse the applicability of the selected guidelines in the current development.
- Develop directly applicable design rules from the selected guidelines.
- Plan and perform frequent accessibility evaluations based on the selected sets of guidelines during the development process.

As mentioned previously, the first step towards developing accessible web applications is to find and select the most adequate sets of guidelines. The most accepted and used sets of guidelines are the WCAG 1.0 [2] and Section 508 [3]. These sets contain general purpose accessibility guidelines. However, for some developments it is necessary to apply also more specific guidelines such as sets of

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 3–10, 2007.
© Springer-Verlag Berlin Heidelberg 2007

guidelines for specific application type (e-learning, tele-working), access device (mobile devices) or user type (elderly, children, blind, deaf). For example, when developing a web application for elderly people specific guidelines as the ones defined in [4], as well as general purpose guidelines should be considered. The following figure, Figure 1, shows the taxonomy of the different types of existing web accessibility sets of guidelines.

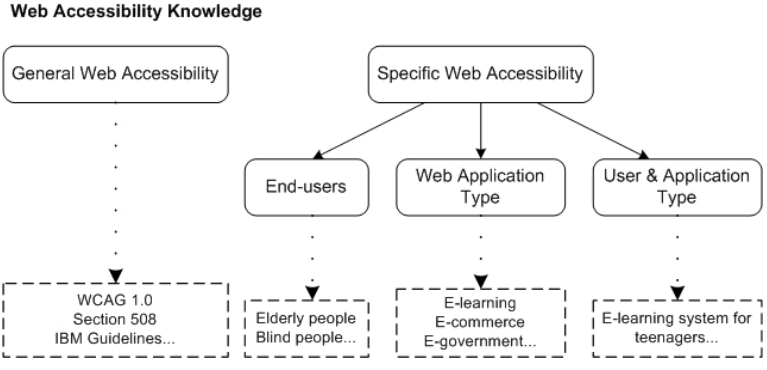


Fig. 1. Taxonomy of web accessibility sets of guidelines

Due to the diversity of formats and structures used for defining accessibility guidelines, finding, selecting, applying and evaluating these guidelines require a great effort from web developers. There are several automatic tools which assist developers evaluating the accessibility of web pages but most of them are based on general purpose sets of guidelines. Therefore, they are not flexible enough to evaluate guidelines for specific application type, user type or access device.

This paper proposes a framework for guidelines definition, edition and sharing which feeds an automatic accessibility evaluation tool. It allows centralizing a guidelines repository from where guidelines can be obtained for customized evaluation purposes. Due to the centralized nature of the repository, sharing guidelines among developers and researchers becomes easy. With the help of this web application, the communication gap between accessibility experts and developers is bridged.

Therefore, this framework is useful during all the development process of accessible web applications since it makes possible to find sets of guidelines according to specific characteristics, defining new guidelines and selecting the sets of guidelines for accessibility evaluations.

2 Related Work

Due to the different formats and level of detail used for defining accessibility guidelines, it is necessary to develop a unified format which contains the required characteristics for specifying guidelines from different sets. The uniform format is the basis for the development of such a management framework. Therefore, it should guarantee that all the necessary data for guidelines evaluation process is included. In this sense, several approaches can be found in the literature.

Vanderdonckt and Bereikdar proposed the Guidelines Definition Language, GDL [5] and recently Leporini et al. the Guidelines Abstraction Language, GAL [6]. These guideline specification languages make possible adapting quite straightforwardly to new guideline versions or novel guidelines.

However, these guidelines specification languages are mostly based on general purpose accessibility sets of guidelines. Consequently, some specific purpose guidelines may not be defined as any previous study of specific accessibility sets of guidelines and their definition in those languages is not provided. In addition, the developed definition languages are quite complex and appropriate tools for defining, editing, sharing and searching for accessibility information are needed.

In one of our previous work [7], UGL a new language for guidelines definition developed based on a comprehensive study of different sets of guidelines is described. This language has been selected for the development of the management framework proposed in this paper.

In respect of guidelines management tools, previous research has been carried out. SIERRA [8] is one of the first approaches for managing usability knowledge by a software tool. However, this tool does not support any evaluation process. Sherlock [9] manages usability guidelines by a client-server system and evaluates automatically only some of the defined guidelines. Another system, called GUIDE, for managing usability guidelines and storing the guidelines applied for a particular application development is presented in [10]. Nevertheless, none of this approaches support completely the development process.

More recent approaches, such as Mariage et al. [11] and Leporini et al. [12] are useful throughout the development process of web applications including the evaluation stage. Both aim at abstracting the interaction with accessibility guidelines with graphical interfaces. Unfortunately, both are standalone applications which have some drawbacks compared with a web application. Moreover, the guideline formats used by these applications are not proven to have been developed based on the results obtained from an analysis of the different types of sets of guidelines. Therefore, some guidelines may not be adequately evaluated or either specified.

3 Guidelines Management Framework

Expert users may prefer to directly specify guidelines in UGL in order to gain efficiency and accuracy. Yet, editing UGL directly may turn out to be a difficult task. Mastering its structure could make novel users reluctant to adopt UGL. Therefore, the proposed framework aims at abstracting the interaction with guidelines sets defined in UGL. The framework consists of a web application which guides the user through all the process of guidelines creation, edition and management. The fact that it is a web application has some advantages over a stand alone application such as:

- It is universally accessible from the browser so there is no need to install software or plug-ins.
- Since the application is WCAG 1.0 compliant users with disabilities are empowered to make their contribution to the guidelines corpus.
- Collected data are stored in the same server where the web application is hosted. The centralization of resources makes feasible access by other users and data are easier to manage.

The framework works jointly with an online accessibility evaluation tool called EvalAccess [13] and is capable of evaluating web pages against the guidelines defined in the framework.

3.1 Characteristics of the Framework

An UGL document contains information regarding web accessibility guidelines in different abstraction levels. The information a level contains is called PACK item which stands for Piece of web ACcessibility Knowledge. *Guideline-set* item contains general information about the guideline set such as the owner, guidelines application environment (general, mobile devices, e-learning...), guidelines end-users (people with disabilities in general , deaf, blind, elderly...), priorities distribution among checkpoints (priority 1/2/3, required/recommended...) and so on. On the other hand, at the lowest level, *evaluation techniques* specify information with regard to mark-up (HTML elements, attributes and their respective values) which is necessary for evaluation purposes. A *Guideline* item is a container which groups *Checkpoints* by their similarity.

As stated in the previous section, the application will lead the user during the process of guidelines creation/edition. The XML-Schema defines the constraints and hierarchical structure of every UGL document: *Guideline-sets* contain *Guidelines* which at the same time contain *Checkpoints* and these contain *Evaluation Techniques*. Therefore, high level PACKs are containers of low level PACKs.

Guidelines creation task is top-down driven and makes possible filling in all the necessary nodes or PACKs in a UGL document. More information about the XML-Schema[1] and which facts led its development can be found at [7]. The following sections describe in detail the functionalities of the management framework.

3.1.1 Guidelines Creation and Edition

First of all the user must sign up in the system and fill in a form regarding to his personal information. Hence, the system has to deal with user management but this task is not a mere personal information collection task, it has to do with the guideline-sets the user creates and shares. Once the user has signed up in the system is able to do the following tasks which are applicable to the following PACKs: *Guideline-sets*, *Guidelines*, *Checkpoints* and *Techniques*.

- Create a new PACK.
- Edit an existing PACK.
- Delete an existing PACK. This implies that all the PACKS that it contains will be also removed.
- View the information of an item. Its title and description are showed together with mark-up information in the case of *evaluation techniques*.

Note that if an *evaluation technique* is going to be defined (as depicted in Figure 2), all the high level PACKs or containers have also to be defined. It is necessary to define the UGL document from the PACK of higher level, in this case the *Guideline Set*, then *Guideline*, *Checkpoint* and finally *Evaluation Technique*. Since this task can be tedious

[1] XML-Schema of UGL: http://sipt07.si.ehu.es/evalaccess3/ugl.xsd. Its representation: http://sipt07.si.ehu.es/evalaccess3/ugl.png

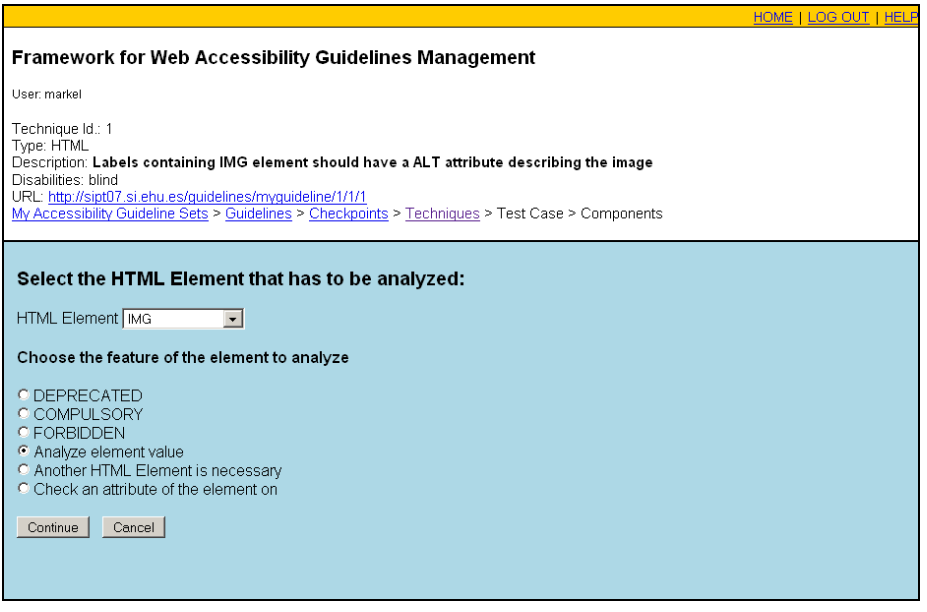


Fig. 2. Element definition in Web Accessibility Guidelines Framework

we are studying the use of predefined UGL templates with the minimum information required so that the evaluation engine will be able to automatically evaluate it.

The process of guidelines definition can be interrupted in any step and the framework will store the state and retrieve it next time the user logs in. Once that a set of guidelines has been completely defined, data are stored in the centralized XML database.

3.1.2 Guidelines Sharing

Once the user has created the PACKs, permissions can be set to them. They can be *not shared*, that is, they will be for personal and private use, *shared* and *shared but not editable* which implies that other users can access, retrieve and use them but the former cannot be changed while the later can be edited. As previously mentioned the system also manages user information such as his personal sets or shared guidelines. Note that permissions of high level PACKs are inherited by lower PACKs.

In order to not to redesign an existing *guideline*, *checkpoint* or *technique*, the system allows users to search for specific information. Search results only show guidelines or techniques which are *shared* or *shared but not editable*. Different search criteria are provided:

- General search: the search will be performed through the accessibility guidelines corpus.
- Specific search: this search is specific to PACKS focused on specific users or applied to particular applications, e.g.: accessibility information (*guidelines*, *checkpoint* and so on) about deaf users or e-learning environments can be obtained.

- Search according to hypertext criteria such as HTML elements and attributes: the users can seek for *evaluation techniques* that contain determined HTML labels or attributes. If they wish they can access the container PACKs such as *Checkpoint* or *Guidelines*.

The following figures show the interfaces for web accessibility guidelines searching (Figure 3) and the results yielded after a query (Figure 4).

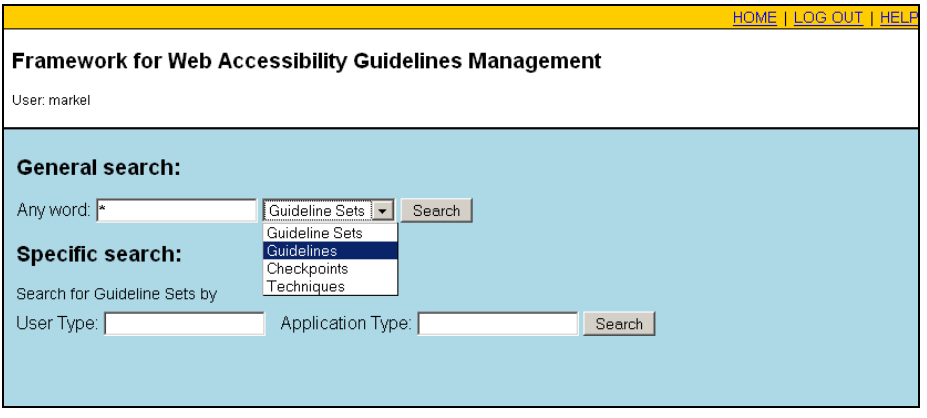


Fig. 3. Interface for guidelines searching

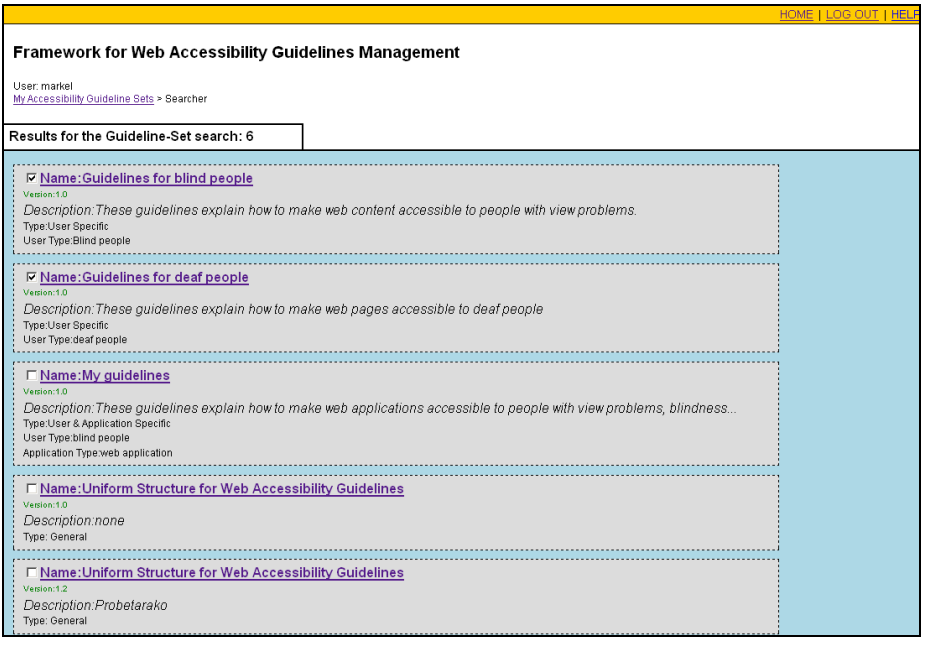


Fig. 4. Results produced after query

Obtained results can be selected and added to a predefined set of guidelines called "My borrowed set" which will keep all guidelines defined and shared by other users. In addition, the PACKs labeled as "shared" could be edited in order to get more personalized accessibility knowledge. Both a copy of the original PACKs as well as the edited version are stored in the XML database in order to keep the consistency since some users may still use the original version (especially the owner).

4 Conclusions and Future Work

The proposed framework assists web developers in creating accessible web applications. It is useful and reliable throughout the development process as different functionalities have been included. In this sense, web developers can edit, update, search for guidelines, include new accessibility guidelines as well as select guidelines for performing automatic accessibility evaluations. Consequently, it is flexible enough to facilitate the development of web applications according to diverse sets of guidelines.

In addition, all the functionalities included in the framework would allow creating a comprehensive repository of accessibility guidelines which could be shared among developers community. Moreover, the implemented language for guidelines definition UGL includes the components for defining diverse types of test cases. Therefore, it guarantees that new guidelines will be easily integrated to the framework.

Acknowledgements. Work of Markel Vigo is funded by the Department of Education, Universities and Research of Basque Government.

References

[1] Abascal, J., Nicolle, C.: Why Inclusive Design Guidelines? In: Abascal, J., Nicolle, C. (eds.) Inclusive Design Guidelines for HCI, Taylor & Francis, London (2001)

[2] Chisholm, W., Vanderheiden, G., Jacobs, I. (eds.): (May 5, 1999) Web Content Accessibility Guidelines 1.0. http://www.w3.org/TR/WAI-WEBCONTENT/

[3] Center for IT Accommodation (CITA) U.S. Section 508 Guidelines http://www.section508.gov

[4] Kurniawan, S., Zaphiris, P.: Research-derived web design guidelines for older people. In: Proceedings of the ACM SIGACCESS Conference on Computers and Accessibility, ASSETS 2005, pp. 129–135 (2005)

[5] Vanderdonckt, J., Bereikdar, A.: Automated Web Evaluation by Guideline Review. Journal of Web Engineering. Rinton Press. vol. 4(2), pp. 102–117 (2005)

[6] Leporini, B., Paternò, F., Scorcia, A.: Flexible tool support for accessibility evaluation. In: Interacting with Computers, vol. 18(5), pp. 869–890. Elsevier, North-Holland, Amsterdam (2006)

[7] Arrue, M., Vigo, M., Abascal, J.: Including Heterogeneous Web Accessibility Guidelines in the Development Process. In: Engineering Interactive Systems 2007. IFIP WG 2.7/13.4, IFIP WG 13.2 and DSVIS Conferences. To be published in LNCS (In press) (2007)

[8] Vanderdonckt, J.: Accessing guidelines information with Sierra. In: Nordby, K., Helmersen, P. H., Gilmore, D. J., Arnesen S. A.(eds.) Proceedings of IFIP TC13 International Conference on Human-Computer Interaction. INTERACT '95, pp. 311–316 (1995)

[9] Grammenos, D., koumianakis, D., Stephanidis, C.: Integrated support for working with guidelines: the Sherlock guideline management system. In: Interacting with Computers, vol. 12(3), pp. 281–311. Elsevier, North-Holland, Amsterdam (2000)

[10] Henninger, S.: A methodology and tools for applying context-specific usability guidelines to interface design. In: Interacting with Computers, vol. 12(3), pp. 225–243. Elsevier, North-Holland, Amsterdam (2000)

[11] Mariage, C., Vanderdonckt: Creating Contextualised Usability Guides for Web Sites Design and Evaluation. In: Jacob, R., et al.: (eds.) Proceedings of the 5th International Conference on Computer-Aided Design of User Interfaces, CADUI 2004, pp. 147–158 (2004)

[12] Leporini, B., Paternò, F., Scorcia, A.: An Environment for Defining and Handling Guidelines for the Web. In: Miesenberger, K., Klaus, J., Zagler, W., Karshmer, A.I. (eds.) ICCHP 2006. LNCS, vol. 4061, pp. 176–183. Springer, Heidelberg (2006)

[13] Abascal, J., Arrue, M., Fajardo, I., Garay, N., Tomás, J.: The use of guidelines to automatically verify Web accessibility. In: Universal Access in the Information Society, vol. 3(1), pp. 71–79. Springer, Heidelberg (2004)

Accessibility Evaluation for Multimedia Content

Chieko Asakawa, Takashi Itoh, Hironobu Takagi, and Hisashi Miyashita

Tokyo Research Laboratory, IBM Japan, 1623-14, Shimo-tsuruma, Yamato, Kanagawa, Japan, 242-8502
{chie,JL03313,takagih,himi}@jp.ibm.com

Abstract. Practices supporting Web accessibility are spread out all over the world. However, the accessibility of multimedia content and rich Internet applications has been little discussed. These types of content are using a lot of visual cues and developing new techniques to make their visual interfaces richer. Therefore, it is generally believed that they are not accessible for people with visually disabilities. However, since there is neither objective data nor tools to measure the accessibility of such content, as a first step to tackle these new accessibility issues we focused on the Flash content. A survey was conducted using five metrics based on the problems faced by screen reader users. We developed a survey tool to assess their severity. Our results showed a complete lack of alternative texts from the developers in the sampled pages. Also, more than half of tested pages were hidden to screen reader users. We conclude that most of existing Flash content is not accessible and there is an emerging need to encourage the use of practices that make Flash content accessible.

Keywords: Flash accessibility, screen reader, GUI accessibility, rich Internet application, blind.

1 Introduction

To date, Web accessibility efforts have mainly been focused on regular static HTML content and awareness has gradually spread among Web designers and developers, with the support of regulations and guidelines. Now, it is time to start focusing on accessibility for multimedia content and rich Internet user experiences. Multimedia content has been becoming increasingly widespread all over the Web. Users can enjoy movies and music animations that are accessed on just like regular HTML pages. In addition, Flash and DHTML are providing rich new user experiences, changing the legacy HTML user experience. The Web is evolving into new dimensions, very different from the static HTML era. These new content types are causing new accessibility problems for people with visual disabilities.

Screen readers read static HTML pages by analyzing the HTML tag structures, and also allow users to navigate through pages using a keyboard without requiring a mouse, even for forms. Such accessible environments are most easily achieved when the HTML pages are created appropriately in accord with Web content accessibility guidelines. Such guidelines do not yet cover accessibility for multimedia content and

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 11–19, 2007.
© Springer-Verlag Berlin Heidelberg 2007

rich Internet applications that are developed using DHTML and Flash. Not has screen reader technology matured enough to support these new Internet technologies. Therefore, it is generally believed that existing multimedia content and rich Internet applications are inaccessible for people with visual disabilities.

Some previous work discussed the difficulties of making the multimedia content and Rich Internet applications (including Flash content) accessible. [1, 2] However, there is neither objective data nor survey results to measure the accessibility of such multimedia content. There were only subjective observations by screen reader users. In addition, there were no tools to evaluate such content's accessibility.

As a next step to tackle the accessibility problems for multimedia and rich Internet applications, we decided to focus on Flash, since it is widespread and has characteristics of both multimedia content and rich Internet applications. We conducted a survey to objectively measure the Flash content accessibility based on metrics for problems faced by users. After deciding on the major problems, a survey tool was developed to assess these metrics quantitatively.

In this paper, we initially discuss the issues faced by screen reader users in accessing Flash and summarize the basics of what Flash developers should do to make this content accessible. Then the survey method is outlined, including the tool design. The survey results are reported and discussed, followed by conclusions and future work.

2 Issues Faced by Screen Reader Users

We categorize Flash content into three types taking accessibility issues into account. They are Flash-based GUIs, Flash vector animation, and Flash movies. Each type has different accessibility issues.

The Flash GUIs shares similar problems with other GUI applications. If screen readers cannot handle a GUI, it means screen reader users cannot access that application. For accessible applications, there are published techniques for making a GUI accessible. For example, all of the operations should be accessible with a keyboard, because a screen reader user cannot use a mouse. All clickable objects without text should be labeled with alternative text, because a screen reader cannot present any information for such clickable objects that lack alternative text. This is similar but more general than the need for alternative text for each image link in HTML. There are various other check items, but these two items are the fundamentals of GUI accessibility enablement. Flash content also needs to use these GUI accessibility techniques. There is a document called "Best practices for accessible Flash design" published by Adobe about making the Flash GUI accessible [3]. If Flash developers follow these best practices, the Flash-based GUI will be accessible to screen readers. Nevertheless, it is generally believed that Flash is hard for screen reader users to access.

Flash content using vector animation poses severe problems for accessibility. It is generally used to present visually attractive and intuitive information, including animations and interactions. The view usually changes dynamically in response to mouse operations and unfolds some kind of story over time. Such visual interfaces

have not been considered in the GUI accessibility techniques. Existing screen reader technology cannot handle such rich user interfaces effectively.

Finally, Flash with movies involves a different type of accessibility problem. The movie content usually includes audio information and visual controls let users control the volume, play or stop the movie, pause it, and skip forward or backwards. These controls use slide bars or buttons and assume a mouse will be used to control them. Screen reader users need to find these controls while listening to two channels of sound, the audio information of the movie and the screen reader output. It is not easy to listen to two sound sources at one time, but existing technology does not allow users to control the two sounds separately. Rather both volumes increase and decrease together.

We overviewed the problems for these three types of Flash content, and as a first step, we focused on the Flash-based GUIs. In the next section, we will explain how existing screen readers read the Flash content and summarize the basics of making the Flash content accessible.

2.1 Fundamental Techniques for Flash Accessibility

Flash has two accessibility functions for making content accessible, with which developers can provide appropriate accessibility information to screen readers. The first one is an "_accProps" property that can add metadata such as alternative text for an image or button. Some Flash authoring tools provide a user interface for seeing and setting _accProps. In Flash Studio, the Accessibility Panel is provided to see _accProps and to set other properties of the objects, such as the "tab index". The tab index property can be used to control the navigation order for screen readers. The order is very important for screen readers to read the text information effectively, though the issues related to the navigation order will not be addressed in this paper.
There are five kinds of information attached to each Flash object with _accProps: the "name" is used for specifying an alternative text; "description" is used for a long description; "silent" is a flag to hide the object itself from MSAA (Microsoft Active Accessibility); "shortcut" assigns a keyboard shortcut; and "forceSimple" is a flag to hide all of the object's children from MSAA.

The second accessibility feature is "_accImpl" for widget metadata such as the role or state of a widget. This metadata provides precise information to screen readers through MSAA and helps screen readers to appropriately verbalize widgets written in ActionScript (AS). In the current Flash Studio environment, developers have to use AS to activate this feature. However, only predefined widgets can be activated for this feature, and custom (user-defined) widgets cannot use it.

2.2 Accessibility Issues

Buttons without Alternative Texts
A screen reader can obtain information to tell if the role of an object is a button through MSAA. Therefore, a screen reader can present a button as "button", but users cannot get any useful information from the word "button" by itself. The _accProps should be used for alternative text for each button. However, users often hear only "button" repeatedly. This indicates that Flash developers are generally unaware of _accProps.

To make this environment more accessible, the newer Flash players automatically search for text information physically near to each button and read that text along with "button" as a description. This is much better than nothing, but appropriate text may not be found corresponding to the actual meaning of the button. Rather than depending on such an unreliable technique, it is strongly recommended that developers use _accProps proactively.

Objects without any reading text
When there is no text in some part of the Flash content, screen readers cannot present any information. Even for an object that is not a button and that does not have any action, _accProps should be used for an alternative text about its role or purpose. If an object is not associated with any text, then a screen reader cannot present any information. If an object is just a graphic with no meaningful information, developers should consider using the "silent" option in _accProps. Care should be taken in this case to be certain that the object can really be ignored completely.

Fragmented text
While navigating through Flash content, users often face situations where the up and down cursor keys only move character by character. Usually, screen reader users can use the four cursor keys to obtain information about the screen, with the up and down keys moving between lines or elements and the left and right keys moving between characters. This is similar to the usual editing functions. However, when all four arrow keys only move between and read characters, then it becomes very hard to understand even one word, and it takes far too much time to get any useful information. If a user needs to press a cursor six times just to hear the word, for example, "l, i, s, t, e and n", they will obviously be discouraged from listening to the page.

The situation happens when a word is divided into separate characters and each character is assigned to an object individually. When using such design methods, the divided word might be animated character by character, so that the content is visually attractive. To permit such design methods while supporting accessibility, the Flash developers could use the "name" and "silent" options. The object corresponding to the first character "l", in the current example, could associate the "name" option with "listen" as alternative text. Then the objects for the other five characters can be associated with "silent". There are various other solutions, but the point is that by using the _accProps options appropriately, the Flash GUI can be much more accessible.

Reporting on the existence of Flash content
When Flash content appears in HTML, screen readers give messages such as "Flash movie start" and "Flash movie ends". If there is any text information to be read, it reads it between those two messages, and users can navigate using the cursor keys as usual. However, sometimes nothing is read except these two messages. This suggests that there is no readable text in the Flash content. Perhaps the Flash content is only an object? Maybe it's only an advertisement? Even in such cases, it is recommended to provide text information such as "advertisement", so that users will not worry about what they missed (or to make them aware of lost information).

Windowless mode option
Using the "windowless" mode creates very critical and severe issues, because once this mode is active, the Flash content becomes hidden from screen readers. There is

no way for blind users even to be aware of its existence. Flash content with the windowless mode sometimes contains very important information such as lists of clickable buttons. It is often used for controlling the visual appearance of the Flash content (in a URL), so it is very difficult to avoid using it if the accessibility problems are to be solved [4].

To solve such problems, the Flash player should not hide the MSAA information from the screen readers. However, currently the windowless mode is hidden, so the only thing Flash developers can do is avoid using the windowless mode for significant content.

3 Survey Method

As described in the previous section, most screen reader users believe Flash content in general is inaccessible. However there is no objective data to support this assumption. Therefore we conducted a survey to objectively measure the accessibility of Flash content.

3.1 Criteria for the Survey

As the criteria for our survey, we defined the following five metrics:

1. The fraction of pages with windowless Flash content and the fraction of Flash content that uses windowless mode.
2. The fraction of buttons without alternative text.
3. The fraction of the Flash space in a page versus the HTML area.
4. The fraction of Flash content with fragmented text.
5. The fraction of Flash content using _accProps.

3.2 Survey Tool and Detailed Method

Figure 1 shows the architecture of the tool for the survey. The Flash content inspector was developed to measure the five metrics. A typical Flash webpage contains multiple Flash player objects in the page. Each Flash player object exposes multiple MSAA objects with the IAccessible interface. Six types of MSAA roles are used by Flash player, the “button”, “graphics”, “label”, “text”, “client-area”, and “animation” roles. Each MSAA object also provides “name” and “area size” information. The tool stores its data into a CSV file used to calculate the final results.

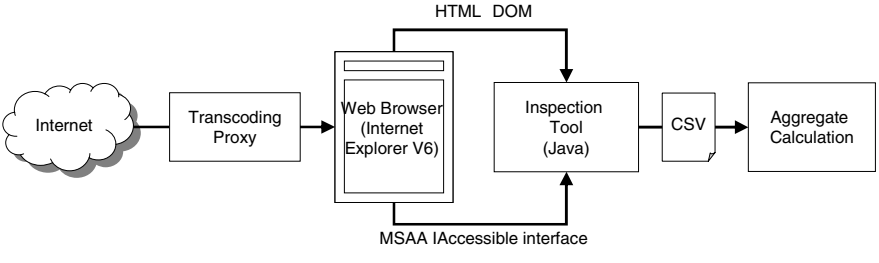


Fig. 1. Architecture of the Survey Tool

The survey tool records the following information for each webpage into the CSV file.

- URL of the tested webpage
- Total number of Flash objects and number of windowless-mode Flash objects in the webpage
- Size of the webpage and the total size of the Flash objects

The survey tool also records the “area”, “role property” and “name property” for each MSAA object in each Flash Player.

The existence of the _accProps and _accImpl properties in a Flash object can not be determined by examining the MSAA object or from the HTML DOM in the Web browser. Therefore our survey tool uses a transcoding proxy to look into the contents of each Flash object [5]. The proxy embeds a component into each Flash object to check for the existence of _accProps or _accImpl. If they exist, the component displays a message on the screen.

The detailed calculations are as follows:

1. **Recognizing windowless mode**
   The survey tool checks “window handle” and “class name” to recognize windowless-mode Flash Player objects. A window-mode Flash object in Internet Explorer exposes an MSAA object, which contains a valid window handle with the class name "MacromediaFlashPlayerActiveX". However a windowless-mode Flash object exposes an MSAA object with the invalid window handle 0.
2. **Counting buttons without alternative texts**
   When an object’s role is button and it has no name information, then it is counted as a button without alternative text.
3. **Calculation of the fraction of Flash content versus HTML**
   An MSAA object with the “client-area” role has a value for its area as a property. The total area of the Flash content can be calculated by summing the areas and subtracting the overlapping areas.
4. **Fragmented text**
   When the “name” property of an object is a single alphabetical character, it is classified as fragmented text.
5. **Checking for the existence of _accProps**
   As described above, the existence of _accProps was checked semi-automatically by the experimenters. The reports about _accProps from the embedded components were recorded.

3.3 Selection of Survey Sites

We followed these rules in selecting candidate websites:

- The websites were from three industrial groups: entertainment, retail, and beverages (selected from the Fortune 500 list).
- The websites were searched for pages containing Flash content, starting from the top pages.
- When there were too few pages with Flash content, then the search went deeper into that website. (Basically a breadth-first search.)
- By repeating this process for the three selected websites, 75 pages were selected for testing.

4 Results

1. **The fractions of pages with windowless Flash content**
   A total of 64 Flash objects out of 127 objects in the 75 pages were windowless (Table 1). A total of 28 pages out of the 75 pages used windowless mode only, which means the Flash content in 37% of the pages was not at all informative for screen reader users. Therefore 63 Flash objects in 47 pages were suitable for this survey. Overall, this shows that the windowless mode is used frequently and screen readers are losing the information in more than half of the Flash objects.
2. **The fraction of buttons without alternative texts**
   Of the 63 Flash objects, 35 of them contained a total of 505 buttons, and 53% (268 buttons) did not have alternative text. Figure 2 shows the distribution of the buttons with alternative text for each Flash object. The distribution is rather bimodal, with unusually large numbers of objects at the two extremes. This indicates that some Flash content has no alternative text for its buttons, while other content provides it for all of the buttons. This suggests that the developers who are aware of alternative text tend to provide it widely, while other developers ignore it completely.
3. **The fraction of Flash space in a page versus HTML area**
   Figure 3 shows the fraction of the Flash space in webpages, showing a trimodal distribution. Two of the peaks are at the extremes and there is a much smaller peak in the middle. This indicates that most Flash use tends to be divided into two cases, one case occupying 100% of a page, the full-Flash pages, and one occupying less than 10%. The small Flash areas often seem to be for advertisements. The hard to interpret peak appears around 50%. These kinds of webpages seem to be using Flash for a main content object or for menus.
4. **The fraction of Flash content with fragmented text**
   Fragmented text only appeared once in our survey data. In that case, the text was fragmented to move the characters independently as a visual effect. This was fewer than our expectations, but it is a severe issue specific to multimedia content.
5. **The fraction of Flash content with _accProps**
   The _accProps were not used at all among the surveyed pages (Table 1). This means that even the Flash content with 100% alternative text is not reliably accessible, since we could not find any _accProps. It seems that all of the alternative text was extracted by the automatic text detection of the Flash Player, as described earlier. This result suggests that the use of _accProps is not well known yet.

Table 1. Classification of Flash objects

Classification		Number of Flash objects
windowless		64
window-mode	without_accProps	63
	with_accProps	0
Total		127

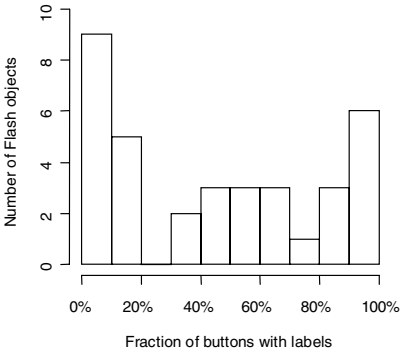


Fig. 2. Distribution for the fraction of buttons without alternative texts

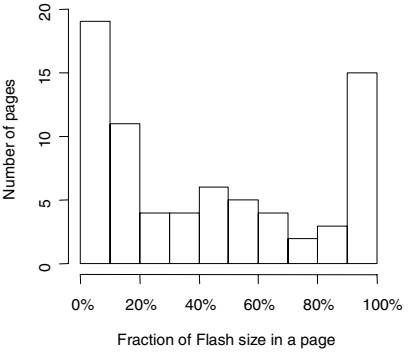


Fig. 3. Distribution for the fraction of the Flash space in a page versus HTML

5 Discussion

The survey result showed how inaccessible the existing Flash content is. We did not find any Flash content using _accProps. This means we could not find any content with explicit consideration of accessibility. Even worse, all of the Flash content in 37% of the evaluated pages was unrecognizable for blind users because of the windowless setting. This was much worse than we had expected.

It would be simplistic to try to blame Adobe for this situation. Adobe has actually made efforts to support Flash accessibility, but the applications of the tools tend to give priority to visual impact, with little or no thought for accessibility. Given the nature of the Flash content and the difficulty of making such visual content accessible, we have to consider the issue more generally, and not dismiss Adobe's sincere efforts. In particular, the automatic detection of button labels is an effective method to automatically provide text alternatives for screen reader users (Figure 2), even though _accProps is being ignored. In many cases, the Flash object itself reveals some of its own information though its audio track.

It is clear that the related technologies should be improved to support more accessible Flash content. Screen readers can be improved for both short term and long term solutions. One example would be notification of windowless content. Screen readers could help screen reader users know that they are losing Flash content when visiting pages with windowless mode content. That way, they would not need to spend time looking for useful information in those essentially inaccessible pages. However, since the Flash content is extremely widespread, we hope the user interface technology in screen readers can improve to support Flash accessibility.

Flash accessibility checking tools are also needed to share knowledge of the existing problems with larger groups of people including auditors, developers, designers, and so on. It would be very useful if more text information is provided by _accProps or automatically extracted by the Flash Player. In order to encourage designers to use _accProps, authoring support tools are necessary.

The best practices should be encouraged, and activities or campaigns in public may help raise awareness of the problems. Some reports have been published, but only a

few and they have had little impact on the developer community. A wide variety of efforts is required to make Flash accessibility activities widespread.

6 Conclusion

Even though it is generally believed that Flash content is inaccessible for screen reader users, there was no objective data or tools to evaluate the situation. Therefore we conducted a survey to objectively measure the accessibility of the existing Flash content. After analyzing the real problems faced by screen reader users, we decided on five metrics for a survey and developed a tool to measure these metrics. A total of 75 existing pages with Flash contents were selected for the survey. The survey results showed that none of tested pages contained any alternative text from the developers. More than half of the Flash objects were hidden from screen readers, because they used the windowless mode.

Without more understanding of the problems and remedial work, Flash will remain inaccessible. This will increase concerns that the Web is becoming inaccessible. After finding solutions for Flash-based GUIs, we want to consider next steps, which are technically more challenging. To date, the accessibility of multimedia content and rich Internet applications is little discussed, even in academia. It is time for us to start addressing these new accessibility issues.

References

1. King, A., Evans, G., Blenkhorn, P.: Webbie: a Web browser for visually impaired people. In: Proceedings of the 2nd Cambridge Workshop on Universal Access and Assistive Technology, pp. 35–44 (2004)
2. Richards, J., Hanson, V.: Web accessibility: a broader view. In: Proceedings of the 13th international conference on World Wide Web, pp. 72–79 (2004)
3. Regan, B.: Best practices for accessible Flash design (2005) http://www.adobe.com/resources/accessibility/best_practices/best_practices_acc_flash.pdf
4. Saito, S., Takagi, H., Asakawa, C.: Transforming flash to xml for accessibility evaluations. In: Proceedings of the 8th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '06), pp. 157–164 (2006)
5. Sullivan, S.: Flash, DHTML Menus and Accessibility http://www.communitymx.com/content/ article.cfm?cid=E5141

Tenuta: Strategies for Providing Guidance on Usability and Accessibility

Nigel Bevan[1], Helen Petrie[1], and Nigel Claridge[2]

[1] Department of Computer Science, University of York, Heslington, York YO10 5DD
[2] Scandinavian Usability Associates, Angantyrvägen 11, Djursholm, Sweden 182 54
tenuta@cs.york.ac.uk

Abstract. The EU Tenuta project was funded to assist EU eTEN projects improve the usability and accessibility of new e-services that are being validated on a European scale. The challenges facing Tenuta included how to initially convince projects that these were important issues, how to gain the confidence of projects that they would benefit Tenuta assistance, and how Tenuta could make an effective contribution with only a few days support available to each project. The eventual strategy was to arrange activities such as training courses where we made personal contacts with projects, which resulted in being invited to provide on-site support. We carried out usability and accessibility evaluations before visiting projects, and reported the results as part of a two to three day workshop with the project, which included assistance with redesign, and a roadmap for future activities. Projects gave very positive feedback on the benefits obtained.

1 Introduction

The European Union (EU) eTEN programme [3] supports the roll-out of e-services across Europe, based on the results of previous EU-funded research projects. EU research projects tend to be technically driven, so usually give little consideration to the usability and accessibility of the e-services. But the eTEN programme is intended to validate use of the e-services in different markets, so is much closer to providing a commercial service, where usability and accessibility should be key issues.

The objective of the EU Tenuta support action [5] is to improve the usability and accessibility of these EU e-services, by providing guidance on how to design systems that are both accessible and usable.

Currently accessibility is primarily concerned with sensory, physical and cognitive compatibilities, while usability is primarily concerned with cognitive compatibility. In most organizations they are the responsibility of people with different skills (technical and human). Neither approach alone is sufficient. Only by integrating these approaches in e-service development can disabled and older be given increased autonomy to play a full role in society.

Many organizations pay lip service to ease-of-use without employing any specific methods to ensure that it is achieved. Many web developers know that web

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 20–27, 2007.
© Springer-Verlag Berlin Heidelberg 2007

accessibility is important, but many do not understand the intricacies of the Web Content Accessibility Guidelines [8] developed by the Web Accessibility Initiative (WAI) of the World Wide Web Consortium and how to apply them.

2 Tenuta Support Action

A paper at HCI International 2005 [1] reported the objectives of the Tenuta project and the challenges that it faced. We anticipated that the main problem would be to provide effective support within the constraints of very limited resources of between 3 and 10 person days of on-site assistance for each project supported.

The eTEN programme management expected that offering free usability and accessibility support would be popular with eTEN projects, but the first announcement of our services by eTEN met with no response. What we gradually learned was that although most projects had little experience of usability and accessibility, they were reluctant to invite an unknown third party to assist them, even one endorsed by the eTEN programme management. Almost all projects that we have now assisted had prior personal contact (e.g. via one of our training courses) or personal recommendation.

So our strategy was to find opportunities to meet project members, by attending meetings and workshops, and organizing training courses. Projects were much more willing to initially commit to sending people to a two or three day training course, than committing to even one or two days on site support. Following attendance at the course, many projects requested more specific support.

2.1 Training Courses

The objective of the two-day training course was to familiarize participants with usability and accessibility issues, and give them some training in basic techniques that they could use themselves.

The accessibility day started by explaining the characteristics and needs of disabled e-service users, followed by a demonstration of use of a screen reader by a blind user navigating a web site. Many participants commented that this gave them a real insight into the problems encountered by blind users.

The course explained the W3C guidelines [8], and gave participants the opportunity to try using tools [7] to check existing web sites for conformance with the guidelines. However, we emphasized that accessibility could only be achieved by working with users to understand how they interact with e-services and by undertaking some user testing.

For the usability day, it was assumed that the audience should already have a good understanding of user requirements for their project, so the emphasis was on how to use this knowledge to produce appropriate designs. The course included an explanation of the business case for user centered design and an overview of how user centered design methods can be used during development.

We had to decide what to focus on for the main content. One approach is to teach heuristic evaluation principles and give participants experience of carrying out heuristic evaluations. This has the advantage of helping developers understand how

basic usability principles can be incorporated into design. But it was felt to be too ambitious for a short course.

We decided instead to teach people how to carry out simple user testing. A step-by-step process was defined for planning and carrying out a user-based evaluation. Emphasis was placed on the need for users to carry out realistic tasks, and for the observer to only provide assistance as a last resort. Simple user testing could potentially enable developers to identify major usability issues with the minimum of training. One objective of this part of the course was to convey the message that although users might be quite satisfied after having seen a product demonstrated, they may find that they are totally unable to use the product on their own.

Most of the participants' implementations were web-based, so we also introduced them to the guidelines for effective web design and usability for information-oriented sites produced by the U.S. Department of Health and Human Services [6]. These guidelines have the advantage that they have been through a thorough review process, are simply described with examples, and can either be purchased as an inexpensive book, or downloaded free as a pdf. Having explained some typical guidelines, we showed how these could be used to evaluate existing web sites. These guidelines should provide a valuable learning resource for participants after the course.

The course has received very positive feedback, although it appears that some participants still thought there should be simple rules they could follow to make a site usable.

2.2 On Site Support

The overall approach to providing support was:

1. An initial telephone conference to review the needs for usability and accessibility, to clarify the intended context of use of the e-service, and to decide how Tenuta could provide the most effective form of support.
2. A review of the usability and accessibility of the e-service, carried out remotely, in advance of a visit.
3. In most cases, a 2-day on-site workshop to give a brief introduction to usability and accessibility, followed by presenting the results of the accessibility and usability reviews and recommendations for design improvement and discussing a road map for continued review.
4. Follow-up activities to support implementation of the recommendations.

2.3 Context of Use

An essential prerequisite for usability and accessibility work is to have a clear understanding of the:

- Intended user groups for the project e-service.
- Reasons that each user group is expected to use the project e-service.
- Context in which this will happen.
- Tasks that each user group will perform.
- Relative importance of the user groups and tasks.
- Importance of supporting users with disabilities.

Although eTEN projects should already have a clear understanding of their potential users, our task is to help projects stand back and evaluate their solutions from the perspective of their users.

2.4 Expert Evaluation

Having gained basic information on the intended context of use from an initial telephone discussion, Tenuta consultants carried out an expert-based evaluation of the usability and accessibility of the project solutions. This was based on established usability principles, web design guidelines (such as [6]), established accessibility principles (the Web Content Accessibility Guidelines [8]) and testing the solution with a range of assistive technologies (particularly screen readers used by blind people, screen magnification programs used by partially sighted people and support software used by dyslexic people). It included both detailed aspects (such as screen design, terminology and support for navigation) and broader task and user support issues: does the project solution support specific user groups in achieving their goals? The effort required (including producing a report) was 1.5-2 person days by an experienced consultant for usability evaluation, and about 3 days by an experienced accessibility consultant to check a sample of 10 pages against the 10 most commonly violated WAI checkpoints [2, 4, 8] and conduct some testing with assistive technologies.

2.5 On-Site Workshop

To maximize the participation by project members, where possible an on-site workshop was arranged. The prime objective of the workshop was to help the project understand how to improve the usability and accessibility of their e-service. It started with a brief introduction to usability and accessibility.

Next, we stepped though the project e-service, discussing the usability and accessibility problems that have been found, and their potential solutions. At this point it is important to gain the support of the project members, so that we can help them implement their own solutions, rather than appear to be imposing a solution on them.

The meetings were most successful when we were working with a group of developers who were in a position to implement the conclusions in a planned new version. We have had most success when our meeting with the project has been hosted by the developers with whom we can work to implement the results of our findings. With one project we only had half a day to present to a project management group who seemed reluctant to accept the implications of the usability findings. The review of usability often raises wider issues of the extent to which the real user needs are understood: why will people use the project e-service, and can use be sustained after the end of eTEN financing?

Feedback that we received early on and incorporated into later workshops included providing a structured methodology for usability testing, and references to further information that could be followed up by projects.

If many cases we went on to assist the project in creating plausible design solutions to the problems that have been identified and agreed.

2.6 Follow-Up Activities

One problem with providing assistance on a consultancy basis is that it is not always clear whether the project has acted on our findings. Sometimes the project management is driven by meeting pre-planned deadlines, and is unwilling to accept the risks they run by proceeding with an unusable solution.

3 Example Case Studies

3.1 eParticipate

eParticipate is creating a European service based on experience in the UK of assisting Local Authorities in the effective use of multimedia and webcasting to create and deliver audiovisual content that really makes an impact in local communities.

To assess the accessibility of the webcasting eService, it was decided that testing by users with disabilities who use assistive technology was vital, because of the use of an embedded media player in the service. Such a service can (and has) been made accessible in other projects we have advised, but to assess the real accessibility of the solution goes far beyond the specificity of the current WCAG accessibility guidelines. The particularly difficult case is whether the solution will be accessible to screenreader users.

Therefore two highly experienced screenreader users were recruited to test the solution, who could make most use of the functionality of their screenreader, and possibly try different methods for accessing the webcasts. If these users could successfully use the solution then it could subsequently be tested by less experienced screenreader users and users of other assistive technologies.

A testing protocol was established by Tenuta in collaboration with the eParticipate team, to ensure that all the key functionalities of the solution would be assessed for accessibility. This took the form of a scenario of five tasks that the users would work through, trying different tasks with the webcasting system.

Two highly experienced users, one who uses JAWS and one who uses WindowEyes were recruited to take part in this initial testing. The JAWS user worked through the scenario, but could not access the webcast at all. Neither could he access some of the information in the titles to the webcast which the team had thought might be accessible. The WindowEyes user had to cancel his session due to family problems, but given the problems by the JAWS user, it was decided to postpone the session with the WindowEyes user until more work has been done on the accessibility of the solution based on the results so far. The JAWS user has offered to re-test the solution remotely when further work has been undertaken.

3.2 EuroWorkSafe

EuroWorkSafe provides information, services and a database of resources related to occupational cancer risks and prevention. The project has carried out a thorough analysis of user requirements, and identified clear scenarios of use for medical, professional and educational users at different sites around Europe.

Tenuta representatives visited the project for two and a half days soon after the first trial version of the web service was available.

Tenuta provided brief training in usability and accessibility, and then reported on the results of a usability and accessibility review of the site. Several changes would be required to the current site to make it accessible to users with disabilities, so it was agreed to help the project ensure that the next version of the site was accessible and then carry out a full accessibility audit.

The main issues revealed by the usability review were that there were some inconsistencies in navigation, that some features of the site are difficult to find and that some technical terminology might not be familiar to the intended users.

Detailed discussion of the users' objectives in coming to the site suggested three main purposes: looking for general information on occupational cancer, searching for specific resources related to occupational cancer, or interest in consultancy services provided by EuroWorkSafe. The EuroWorkSafe project proposed a new navigation structure to more directly support these tasks, which also resulted in a more consistent approach to navigation. This will be implemented in the next release.

Tenuta also recommended that EuroWorkSafe observes three to five representative users using the new site, individually carrying out typical tasks without any assistance. Similar user testing of the current site could also help inform the new design.

3.3 Medical Web Site

Tenuta reviewed the usability of a web service that provides information for doctors that is not easily available elsewhere. The project had demonstrated their web solution to doctors and incorporated some feedback. The project was visited by Tenuta shortly before a planned validation in several countries by 500 users.

The usability review suggested that it is unlikely that any experienced web user who makes the effort to familiarise themselves with the site will encounter any major problems. However, doctors are often busy and impatient, and may not devote the effort to become familiar with the site, and are unlikely to spend much time reading documentation. Users who have initial problems are unlikely to return. So the site should be as easy as possible for first time users.

In addition, the acceptability and usefulness of the site for doctors may depend on minor issues that can only be identified when doctors try using the site, so Tenuta strongly recommended initial in depth pilot testing with a few selected doctors.

The project had also prepared a very complex validation questionnaire. Tenuta recommended that it should be limited to about 10-15 simple questions, as additional questions or a complex questionnaire would significantly reduce the response rate.

Following the visit, Tenuta contacted a doctor by telephone, explained what the site was for, and asked them how they might use it. The doctor was invited to try carrying out their own tasks, and to explain their experiences over the phone. After only a few minutes the doctor commented: "It is so long winded that I would never in a million years use it". There were problems with the complexity of the interface, interpreting the results, and incomplete information in the underlying database. If these findings were replicated with other users, it would confirm the danger that

validation trials using the existing version of the web site could be a complete disaster!

This highlights the risk of just relying on demonstrating a solution. The way the web site works may make sense to the audience at the time, but nevertheless leave them mystified if they sit down and try to use it unaided.

4 Conclusions

Although we have been very successful in raising awareness in the organizations we have worked with, it is difficult to know to what extent our original objective of transferring skills to the projects has been successful. The challenge for future support activities is to find ways to help projects build usability and accessibility into their development processes at an earlier stage. This mirrors a similar challenge to consultants working in industry.

In almost every case we have found that projects have very little awareness of the importance of issues associated with usability and accessibility. In future, we expect the EU to be more demanding when vetting proposals for funding, particularly for near-market projects such as those in the eTEN programme. For example:

- Requirements for usability need to be derived from an in-depth understanding of user needs. As there may be a high risk of product failure in the absence of these requirements, more emphasis should be given to project competence in defining usability requirements, and designing systems that meet these requirements.
- We encountered the common misconception that the best way to get feedback on a service is to demonstrate it to potential users. Projects working with potential users should be required to include one to one usability testing as part of the validation activities.
- There is increasing emphasis within the EU on producing products and services that are accessible to users with disabilities and older users. Legislation already exists in some member states, and a common EU position is likely in the future. In these circumstances one would expect projects to be required to take reasonable steps to develop systems that are accessible.

We did not expect that it would be easy to provide effective support to EU projects, but even we were surprised how difficult it was to gain the initial interest of the projects. Ironically, after we had been invited to provide some support, several projects suggested that Tenuta should be more involved in the projects from their beginning, giving continuous support throughout the lifecycle of the project.

What we discovered working with projects was rather what we expected: very little consideration had been given to how to make project e-services usable and accessible.

But we have had consistently positive feedback on the benefits provided by our support:

- “A new insight into the user experience has been provided.”
- “Enabled us to revise the portal with new layout and improved interaction mechanism.”

- "The usability evaluation has been very helpful in discovering new functionalities that should be supported by the software products."
- "Focused the work on the most critical aspects, in order to reach the project objectives."

As this could have a major impact on the eventual commercial success of the planned services, it is important for the EU to have more stringent criteria for competency in this area as part of the conditions of their funding.

References

1. Bevan, N., Claridge, N., Petrie, H.: Tenuta: Simplified Guidance for Usability and Accessibility. Human Computer International 2005. In: Proceedings HCI International 2005, Lawrence Erlbaum, Mahwah, NJ (2005)
2. Disability Rights Commission: The Web: Access and inclusion for disabled people. London: The Stationery Office (2004) Available at: http://www.drc-gb.org/library/formal_investigation_report_w.aspx
3. eTEN programme: http://ec.europa.eu/information_society/activities/eten
4. Petrie, H., King, N., Hamilton, F.: Accessibility of museum, library and archive websites: the MLA audit. London: Museums, Libraries and Archives Council (2005) Available at: http://www.mla.gov.uk/resources/assets//M/mla_web_accessibility_pdf_6541.pdf
5. Tenuta support action. Available at: http://www.etenuta.org
6. U.S Department of Health and Human Sciences: Research-Based Web Design & Usability Guidelines (2006) Available at: http://www.usability.gov/guidelines/
7. Vision Australia: Web Accessibility Toolbar. Available at: http://www.visionaustralia.org.au/ais/ toolbar/
8. W3C, Web Content Accessibility Guidelines 1.0 (1999) Available at: http://www.w3.org/TR/WCAG10/

Remote Navigability Testing Using Data Gathering Agents

Martin Gonzalez-Rodriguez, Fidel Diez-Diaz, and Agueda Vidau

The Human Communication and Interaction Research Group (HCI-RG)
Department of Computer Science, University of Oviedo, Calvo Sotelo, s/n,
33007 OVIEDO – Spain
www.hci.uniovi.es

Abstract. Testing for web navigability is one of the most important steps in the design of web product. However, navigability testing is an expensive technique which requires the use of many human resources. The user behavior recorded in such kind of tests is usually affected by the artificial situation created in the laboratory. In order to avoid such kind of drawbacks, we have developed ANTS (Automatic Navigability Testing System), an application able to observe and to record spontaneous user behavior in his/her own home.

Keywords: Agents, Navigability Testing, Usability Testing, Remote Testing, ANTS.

1 Introduction

We are going to introduce a new system developed to help designers to build up web sites more usable. Before doing so, we are going to introduce basic concepts of usability and how this quality feature of web site can be measured. Then we will evaluate several different tools developed to measure usability, evaluating their advantages and disadvantages.

2 The Usability Lab

The success of any web site depends on the users who are going to use it in the future. Those users should find the site satisfactory, effective and efficient (ISO Ergonomic Requirements, ISO 9241 part 11: *Guidance on usability specifications and measures*) [1]. If the user feels uncomfortable while surfing the web, it would be very easy for him/her to close the browser and look for another web site. For a business, this event would mean a high waste of money as well as allowing rival companies to catch those clients.

If we want to know how the user's navigation performs, we have to make use of a usability laboratory. The usability lab offers powerful and precise tools to obtain tons of data which helps designers to analyze the weak points of their site: difficulty of use, misunderstanding of key features the interface, incompatibilities between the user's and designer's mental model and, of course, to user's interaction.

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 28–33, 2007.
© Springer-Verlag Berlin Heidelberg 2007

However, the use of physical usability laboratories has also a series of important drawbacks:

- The evaluation of the navigation requires a high precision and constant observation of the volunteers. The observation time is usually very long, so a lot of resources must be spent.
- Due to the high prices of the evaluation, the number of volunteers in usually small. Therefore, the accuracy of the results obtained decrease.
- Testers have to choose volunteers among people with the same cultural level of the intended target user, hopefully finding recruiting users whose behavior matches that of the average user, so it is difficult to discover new types of users no already identified.
- Volunteers know that they are being observed. This fact adds external noise to the evaluation process. Nervousness, confusion, and similar reactions are usually recorded by the usability laboratory video cameras.
- Knowledge about specific details of the experiment also influences the user's behavior when performing navigation tasks. Volunteers know what it is expected from them, so they may adapt their own techniques of navigation to the model of navigation supplied by testers (designer's mental model).
- Usability evaluation makes use of computers that match with some specific technical specifications needed to run the application. The awful truth is that performance of the user's computers is usually far away from the performance of the laboratory computers.
- Since the evaluation of the usability takes place in a laboratory, testers usually miss how the user's computational environment affects user's behavior and performance.

In order to avoid the problems commented, we propose the use of a Remote Navigability Testing approach, replacing human observers by software data-gathering agents. An automatic data gathering process means a cheaper testing process, since the high costs derived from navigability testing are generally caused by hiring a laboratory and human resources for long times of observation.

Some of the advantages of this approach are:

- Users ignore the nature of the evaluation, as well as the role they are playing in it. They only interact with the system under testing.
- Users are not under pressure. They feel comfortable to explore the navigational map provide. External factors do not affect the test results.
- It is possible to obtain information about how the performance of the computer affects how users explore the site.

When the remote evaluation is carried out, the prototypes can be freely distributed as beta versions to be evaluated by all users interested in the application.

3 State of the Art in Remote Testing

Next, we include the main features of some of the most important data gathering tools applied to navigability testing.

3.1 Ergolight

Ergolight [2] offers the Ergolight Interactive Site Usability Reports to evaluate a web site. That product allows the designer to obtain reports about the design problems that affect the user's experience. The report includes a list of pages sorted by cost of effectively. For each page, the reports also include a list of problematic links. To use this service, the designer needs a visit log for the site. Also, the domain is needed in order to allow the inclusion of links between the reports obtained and the analyzed pages. That log file should be created by the server automatically. The analysis report includes the following items:

- Site ranking. Evaluating the ease of use for the web access, and information quality.
- Site diagnostic. It evaluates whether the pages are hard to find, if they are going to irritate the users, if they are going to be displayed slowly, and so on.
- Statistics. It will shows statistics regarding the number of visits, , the percentage of inputs and outputs, the size of the downloads realized, the number of access to each page sorting by the size of them, and so on.

3.1.1 Advantages

The number of variables analyzed regarding the user interaction is high. To see the results, we only have to access the result page. This page is always updated. It brings suggestions to improve those items that should be fixed. Each page is important. Concrete actions like downloads are important. The results include links to specific pages.

3.1.2 Disadvantages

To use this service we have to know how the web server works. This program doesn't obtain information directly from the users. It obtains it and classifies it from the information included into the log archive of the web server. Since the results are mere statistical reports, if we need to increase the accuracy of the results obtained, it is necessary to increase number of visits to our web site.

3.2 WebMetrics NIST

WebMetrics NIST [3] offers a group of products to facilitate the task of determinate whether a web site is usable. This group is composed by:

- WebSAT (Web Static Analysed Tool). It indicates whether the HTML code carries out with the IEEE specifications or with the WebSAT usability guides.
- WebCAT (Web Category Analysis Tool). We can create different kinds of categories and different kinds of objects. Each object belongs to a category. During the test the users have to determinate which is the better category for a concrete object.
- WebVIP (Web Variable Instrumented Program). This product allows the insertion code into the pages to make automatic advice of how the user can do some tasks.

3.2.1 Advantages

The tools are orientated to make easier the web design, not to evaluate the web after it was developed. When the user is asked to perform any action, it is not required to learn new especial procedures.

3.2.2 Disadvantages

The division of the available features into different tools becomes a disadvantage when we it is required more than one tool for the evaluation. WebMetrics uses its own special format (called FLUD, the collaboration between different applications becomes complex.

4 ANTS

We have designed an agent-based tool called ANTS (Automatic Navigability Testing System) in order to provide support for remote testing. The design of ANTS is based on a client-server paradigm, using a design metaphor inspired in colonies of ants. Our system uses a large collection of different ants, which were specially designed for the purpose of gathering specific kind of data. For instance, our system uses text agents to determine user's typing skills, motion agents to measure the user's motion accuracy and so on.

Whenever web document is downloaded, the ant included in the web page is downloaded from the anthill (the ANTS server). When the document is displayed to the user, the ant establishes a communication channel with the anthill, which is used to send the data collected. An ant-assistant object located in the server-side (see figure 1) maintains the communication channel alive during the interactive dialog's cycle of life, collecting the information sent by the ant, classifying it, and sending it to the warehouse object commissioned to store it in the proper user model.

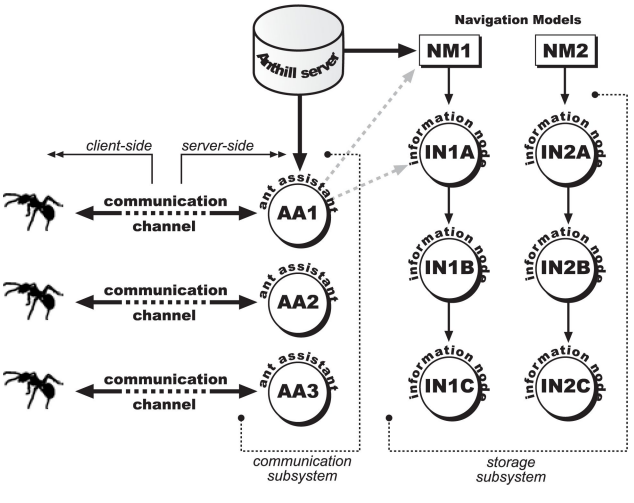


Fig. 1. ANTS design

The ants keep track of every action performed by the user in every user process, generating an abstract navigation model similar to those employed by the Object Oriented Hypermedia Design Model (OOHDM) [4]. This abstract navigation model is used to determine the landmarks, routes and mental models [5] employed by the users when they make use of user process of the application. This information is later used by the testers to determine the user's expertise degree in the use of the application [6].

Specialized agents increase the amount of gathered information or complement it for specific usability tests. When using standard ants we get valuable information about the navigation behavior of a huge amount of users but not for all. For instance, standard ants can get information about the links employed by the user when performing a mouse-based navigation but no information is retrieved when using a key-based navigation.

Visually handicapped users and any other users with similar special interaction requirements need to perform key-based navigation. In this case, the tab key is used to select the required link moving the focus from one link to the next one.

Obtaining information about the key pressed by the users along their navigation session is crucial in this kind of tests, so the standard ants are complemented by text-ants which are able to record any single key pressed by the user. This data is later used by the researchers in order to reproduce the navigation session for an offline analysis.

Mouse-based and key-based navigation can be analyzed separately in order to determine different navigation speeds, level of expertise and interaction requirements in both approaches.

ANTS also includes other agents specialized in gathering specific information to obtain a better view of the interaction requirements of the users. For instance, the motion-ant is able to obtain a mouse motion accuracy rating for each user of the project. This information is quite valuable in order to design the visual appearance of the web. For instance, if the average users are elderly people or any other kind of users with low motion accuracy, the size of the target objects in mouse pointing tasks (links, buttons, etc.) should be increased accordingly in order to make this task easier.

5 Conclusions

When using automatic remote testing tools such ANTS, data flows freely from the origin to the server storage system, where it can be analyzed off-line. As navigation takes place in the own user-computing environment, there is no need to assign expensive laboratory resources for testing.

Testers are free from the boring task of capturing data, so they can focus their efforts in analyzing the results obtained and improving the quality of the navigation maps. As there is no need to assign human resources for conducting user observation, the whole process of usability testing is cheaper.

We want to remark however, that remote testing does not pretend to be a substitution of classic usability testing. Our intention is to optimize usability testing for certain special tasks. Remote testing is able to detect actions but it is unable to observe how those actions are performed. For example, with remote testing it is quite easy to get the number of clicks on a button during a navigation session, but it is

impossible to know the difficulties that the user could have when he/she performed the click. That is a task where classic usability testing seems to be unbeatable.

Acknowledgments. This technology has been developed through projects IB05-152 SEELE: Automatic generation of Web Interfaces dynamically tailored to the requirements of different kinds of disabilities and PC-04-31 GeoFront: Real Time Collaborative Knowledge Manager System for Virtual Worlds. Funds provided by the Asturian Office of Science and Technology (FICYT).

References

1. Ergonomics of Human System Interaction – ISO 9241 (2007) http://en.wikipedia.org/wiki/ISO_9241
2. Ergolight (2007) http://www.ergolight-sw.com
3. WebMetrics, N.I.S.T (2007) http://zing.ncsl.nist.gov/WebTools
4. SCHWABE, D., ROSSI, G.: Abstraction, Composition and Lay-Out Definition Mechanism in OOHDM. In: Proceedings of the ACM Workshop on Effective Abstractions in Multimedia. San Francisco, California (1995)
5. JUL, S., FURNAS, G.W: Navigation in Electronic Worlds. In: a CHI'97 Workshop. SIGCHI Bulletin, vol. 29(4), ACM Press, New York (October 1997)
6. Keith, K.S., PERRY, T.J., KROLCZYK, M. J.: a CHI 97 Workshop. ACM SIGCHI Bulletin vol. 29(4) (October 1997) Michalewicz, Z.: Genetic Algorithms + Data Structures = Evolution Programs. 3rd edn. Springer-Verlag, Berlin Heidelberg New York (1996) http://www.acm.org/sigchi/bulletin/1997.4/karn.htm

Enabling Accessibility and Enhancing Web Experience: Ordering Search Results According to User Needs

Alice Good and Jenny Jerrams-Smith

Department of Information Systems and Computer Applications
Burnaby Terrace, University of Portsmouth, UK
alice.good@port.ac.uk, jenny.jerrams-smith@port.ac.uk

Abstract. This paper presents the results of an exploratory study, which aimed to devise a means of improving web accessibility for users with impairments. It looks at how web pages can be rated for accessibility using specific algorithms. These algorithms analyse and select web pages according to user preferences. The study gathered user defined ratings of web page elements, for different categories of special needs' users that affect accessibility. The algorithms were then created using the results of these ratings. The paper also demonstrates the work in progress in relation to the extension of the Computer Aided Internet Navigation (CAIN) system, which aims to improve accessibility to web based information by re-ordering search results according to user needs.

Keywords: Accessibility; special needs; web navigation; algorithms.

1 Introduction

The purpose of this study is to focus upon addressing the issue of poor accessibility, which affects people with special needs. These needs explicitly relate to people with limitations that affect the ease in which they are able to access web based information, including older adults with age related impairments. The limitations considered here include all types of visual and motor restrictions as well as dyslexia. The scope of this study was focused on the more common impairments and hence other cognitive and language impairments will not be considered.

The strive towards accessibility has been a long-standing, major concern for representatives of special user groups and recent legislation has contributed to the importance of ensuring that these users are enabled and indeed have a right for equal access to information. Buzzwords such as 'Inclusive design' and 'universal access' are increasingly becoming omnipresent in the world of web design, along with numerous sets of guidelines and tools to assist and advise designers in ensuring that potential users are not excluded. As Nielson states: *"It has gotten much easier to advise people on making it possible for users with disabilities to use a website: just follow the official Web Accessibility Initiative Standard (WAI) from the World Wide Web Consortium (W3C)."* (Nielson, 1999, Jakob Nielson's Alertbox). The development of assistive and adaptive software have certainly provided better access to information for this user group though, in terms of equal access, there is no comparison to non-disabled users (Stephanidis, 2001; WAI, 2003; DRC, 2004). There

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 34–44, 2007.
© Springer-Verlag Berlin Heidelberg 2007

is a general consensus that poor design and incorrect coding are key factors in the cause of problems related to accessibility (Nielson, W3C Schneiderman,).Much of the problem is also said to be a result of a lack of consideration to the entirety of user needs being incorporated in the design process; poor design then leads to poor usability (Nielson, 2003; Schneiderman,1999; Preece et al, 2000) The ever-increasing numbers of elderly people contribute to the scale of the problem given their susceptibility to similar disabilities such as reduced vision and mobility impairments. The problem of inaccessibility is not an isolated issue (Eklund & Ziegler 1996; Marshall 1995), other related or resulting factors include digital isolation, social exclusion and disempowerment (NTIA, 1999).

In addition to its renowned value to facilitate learning, the Internet is an expansive tool for reducing digital and social isolation and empowering individuals with disabilities. This then makes it all the more important to ensure these people have access to the information that they require. Difficulties in accessing information reduces the Internet's learning potential and prevents people from participating in the social, political and economic issues that are increasingly being enabled via Information and Communication Technologies. (ICTs) (Carver, 2000). This then increases the likelihood of social exclusion to the disadvantaged. (Schneiderman, 2000) Both disabled and elderly people commonly experience social isolation. This is then another area where the Internet could be particularly beneficial in assisting the maintaining of social contact. Then there is also the subject of disempowerment, the result of necessary reliance upon the assistance of others, that many people within this user group are subjected to. Those that are blind potentially have less need to be reliant upon others for assistance with tasks such as shopping, studying or working. People with mobility and hearing problems could access documentation online instead of facing difficult journeys and problems with communication. Furthermore, the ability to become more independent is enabled, thus increasing the individual's self esteem. Learning potential, inclusion and empowerment are all issues that could possibly go some way to compensate for the restrictions that are imposed by the effects that their impairments have upon their lives (Kobsa & Stephandis, 1998). In short, the improvement of accessibility will not only increase learning potential but is also likely to have a positive effect upon other aspects of their lives, which are affected by the nature of their disabilities. These problems and the issue of accessibility in general is also relevant to older adults.

The goal of universal accessibility has been the driving force behind new techniques of information adaptation to cater for specific needs (Preece et al, 2001; Nielsen, 2001; Scheiderman, 2000). Although there is a wide variety of assistive technology to enable web pages to be made more accessible, this does rely upon correctly coded pages. For example, where pages do not have alternative text for images, screen readers cannot describe a picture to a blind user. In addition, there are elements of web pages that do not lend themselves to modification. For example, the layout of information can present problems for dyslexic users, whilst Web pages with interaction that requires precise mouse movements would present problems for motor impaired users. Users can then find themselves sifting through numerous pages to locate required content. This can be tiresome for even accomplished users but for people with special needs this task becomes even more arduous and frustrating, depending upon the extent of impairment. The content of a web page may match the

user's goal but if it cannot be accessed as a result of poor design, then the content becomes useless. Selecting the presentation order of web pages according to specific needs could then be a possible alternative.

The specific requirements for the intended user group need to be identified as well as working out what adaptivity is needed. Elements of web pages that affect the accessibility for users with disabilities will need to be identified. This will help to understand how information can best be selected and ranked. To enable the pre-selecting of pages according to the user's specific requirements, some means of developing new attributes of web pages, to be stored in the web model, needs be explored. The web model will enable the process of automating the ratings of web pages, according to the specified attributes contained within the user model. For example, the *suitability* of web pages may be categorically rated as: very suitable, medium or unsuitable. Further explanation of this might be where a page's elements are all accessible, then it can be said to be suitable. A few non-accessible items may constitute medium, more than this would then deem the page as unsuitable.

Weightings of web page elements that affect accessibility for visually and motor impaired, as well as dyslexic users needs to be established. User defined ratings of these elements that affect accessibility would enable a set of user preferences to be created. Each category of impairments will have its own set of preferences, and an algorithm would then be used to rate web pages (Good, 2004). A default set of ratings for each category can then be adjusted according to user preference. The WAI gives a comprehensive description of which elements relate to which type of impairment and how they affect the user (W3C/WAI, 2004); some of which have up to nine specified elements. There are numerous elements that can affect a user with a particular impairment, however there appears to be little research conducted on the weightings of these elements even though they can differ in the extent to which they affect accessibility. It is the weightings of these elements that remain the focus of this study.

It is anticipated that to create an algorithm on the weightings of all elements that affect accessibility, would considerably increase the scope of this research. An important consideration has to be the time taken to analyse individual pages. The greater the number of elements, the longer the parsing time will take. In addition, to facilitate the selection of web pages based upon the entirety of specified elements would then significantly reduce the number of accessible pages and would impact upon the user's search criteria being fulfilled. It is also hoped that by limiting the algorithms to three elements, this will enable a preliminary analysis of the feasibility of selecting web pages according to user needs. The purpose of this study is then to gain user defined ratings of three elements for each category of users. Another reason for specifying that the users select only three is to reduce the likelihood of eliminating elements purely upon preference only, although it is hoped that further research will enable this facility to be implemented. It should be noted that this is an exploratory study and as such, the results will be used as an indication of preferences.

2 Method

The aim of this study was to gain user defined ratings of web page elements that affect accessibility. A user-centred approach was adopted for the study with a

combined thinking aloud protocol and an observation study, with a follow-up questionnaire.

2.1 Procedure

Participants were asked to browse some information from the Web relating to news, health information and communication/recreational sites. They were given three tasks:

1. Find out what the current situation on SARS is (Severe Acute Respiratory Syndrome).
2. Find information on how to prevent asthma using the NHS web site.
3. Find a discussion group for gardening

Participants undertook various search related tasks. They were requested to talk through the process expressing their thoughts on the web pages they visited. In particular, they were to voice any problems relating to ease of usage and accessibility of web pages. Observations and comments were noted. After completing the task, participants were asked to specify which elements of the web sites they used affected the ease with which they were able to access the information. They were then asked to select the three most prominent elements and rate them between one and three with one being the element that made accessibility the most difficult.

3 Results

The study yielded two types of data, qualitative and quantitative, which were analysed differently but will be used for comparison. The elements that were rated by users were analysed using a quantitative approach (Good, 2005). The thinking aloud protocol/observation method yielded qualitative data that was analysed using content analysis. It is the user defined ratings of elements that this study is predominantly concerned with and as such, the qualitative data will be used for reference only. It should be noted that there was little difference between the two sets of results. Statistical analysis was not deemed appropriate to analyse the results from this study because of the relatively small sample size and means of selection. It has previously been noted that this was an exploratory study and as such, the results will be used to give an indication of preferences. The results are therefore to be viewed as a pre-requisite for further research into user-defined ratings of elements that affect accessibility.

3.1 User Defined Elements

The participants were asked to list three elements and rate them in the order in which they affected web page accessibility. The elements were given scores according to their ranking. The first element was assigned a score of three, the second element received a score of two and the third element received a score of one.

Blind Participants. All six participants rated the lack of alternative text on graphics being the most significant accessibility issue for them. Three participants assigned

frames as being the second element, with a rating of two that caused accessibility problems with the other three assigning this rating to pop-ups. The third element to cause accessibility issues was text embedded in pictures for three participants and pop-ups for the other three.

Other Visually Impaired Participants. There were fifteen participants who had varying degrees of visual impairment but were not medically defined as blind. Font size caused the most accessibility problems followed by low contrast between the font and the background colour, and with pop-ups being rated third. Other elements that caused problems were: patterned backgrounds; scrollbars; layout; embedded links and blue links. Only one element from the WAI list (W3C/WAI, 2004), was not specified and this was image text that cannot be re-wrapped. With the exception of font size, contrast between background and foreground, and layout, the remaining user defined elements do not feature in the list provided by the WAI. In terms of the second two elements: background colour and contrast between background and foreground, these two could be said to imply the same thing. Background colour can affect readability. Some justification for why the remaining elements were defined by the participants but were not specified by the WAI could be that these elements refer to the user's experience as opposed to accessibility. Results from the observational study did not highlight any additional elements.

Dyslexic Participants. Small font size and pages with large blocks of text were stated as being the elements that caused dyslexics the most problems. Participants rated these two elements equally. Web pages with white backgrounds were also problematic as was inconsistent layout, the lack of pictures and sentence length. The WAI have specified four additional elements that cause accessibility problems for the cognitive/language category. A likely explanation as to why they weren't rated by the participants is because the category includes a number of impairments including epilepsy and thus, video/audio frequencies triggering seizures would not be applicable for dyslexics. Conversely, there were a number of elements defined by the participants, which were not outlined as being problematic by the WAI. These include font size, too much text, background colour and sentence length. The observational study noted that when background colour was mentioned as being problematic, it was often because the colour was white. Other elements noted were lack of headings and very long web pages.

Motor Impaired Participants. There were ten motor impaired participants that took part in the study. The results show small links as affecting accessibility the most for this particular user group. Scroll bars came second with contrast between fore and background colour being third. All the elements defined by the participants were also outlined by the WAI. Conversely, the participants did not mention the following 3 elements as specified by the WAI: time limited response on web pages, no keyboard alternatives and forms that cannot be tabbed through. A likely explanation for this could be that the websites the participants visited did not feature these elements. The observational study highlighted the same elements specified by the participants with the exception of one, that being the lack of search boxes on large websites

4 Discussion

Where assistive and adaptive software have certainly provided better access to information for special needs users, in terms of equality of access, there is no comparison to non-disabled users. The problem itself is widely recognised with numerous interest groups researching new techniques and introducing legislation that will provide better access to future web-based information. However, despite the availability of standards and guidelines there is still a significant shortfall of compliance. It is this shortfall that prevents people with disabilities equal access of information. Where non-compliance to design standards prevails, there will always be users who are faced with barriers. Learning potential, inclusion and empowerment are all issues that are affected by inaccessible web pages. Selecting web pages according to their accessibility could help to reduce the amount of inaccessible information presented to the user and thus increase the value of the World Wide Web.

Table 1. A Summary of the Results

	Blind	**Short-sighted**	**Dyslexic**	**Motor**
1	Images-no alt text	Font size	Font size	Small links
2	Frames	Contrast	Layout	Scroll bars
3	Tables	Pop-ups	Background colour	Contrast

This paper presented the results of an exploratory study to gain user defined ratings of elements that affect accessibility for users with special needs (Table 1). The study included participants with a wide range of sensory, physical and cognitive impairments. Participants were asked to specify and rate, with a scoring of one to three, the elements they felt affected accessibility of web pages. Many of the elements that were rated were ones that had been previously specified by the WAI. This is certainly true for all visually impaired participants. In terms of the top three rated elements (see Table 1), almost all dyslexic participants specified that font size was a problem and over half said that when the background colour was white, it made readability particularly difficult. These two elements were rated one and three respectively, however the WAI does not specify either of the elements as likely barriers. Those participants that experienced difficulty reading black text off a white background used coloured filters or wore precision tinted glasses, with individually chosen prisms and colour when reading web pages. Some of the elements that received lower weightings were also not ones that had been specified by The WAI. A possible explanation for this could be that they were elements that affected the user's experience as opposed to the effect they had upon accessibility. This would seem particularly relevant where over half of motor impaired participants specified that low contrast between foreground and background colours presented difficulties. A possible explanation as to why those elements that were outlined by the WAI were not defined by the participants could be because the web sites explored within the study did not feature them. Another reason could be that due to the small sample size and as such, the participants may not have experienced problems in relation to those elements.

5 Algorithms

The results from this study were used to create a set of algorithms, customised according to user needs. The purpose of the algorithms will be to analyse web pages and assess their suitability for users with specific needs. The algorithms calculate web page accessibility by looking for specific elements known to known to affect a particular type of user. The elements incorporated into the algorithm are the highest three rated by users from the reported study. The following sets of algorithms demonstrate how web page accessibility is calculated.

5.1 Algorithm for Short-Sighted Users

The following rule script shows how an algorithm makes decisions upon selecting the most accessible web pages for a short-sighted user:

```
If (user = 'Short-sighted')
{set page accessibility to 100
        {Check font size
                if (font size is fixed)
                        then reduce accessibility by 20
                if (font size of body text is less than 10pt)
                        then reduce accessibility by 20
                if (font size of body text is less that 12pt but greater than 10pt)
                        then reduce accessibility by 10          }
        { Check contrast between background and font colour
                if (contrast is less than 100)
                        then reduce accessibility by 50
                if (contrast is less than 200 but greater than 100)
                        then reduce accessibility by 40
                if (contrast is less than 300 but greater than 200)
                        then reduce accessibility by 30
                if (contrast is less than 400 but greater than 300)
                        then reduce accessibility by 20
                if (contrast is less than 500 but greater than 400)
                        then reduce accessibility by 10}}
```

5.2 Algorithm for Blind Users

```
If (user = 'Blind')
Total_area_ image = total area of images
na_image = non-accessible image
space_na_image = % of total images that non-accessible image(s) uses // non accessible
image  - no alt text
{       set page accessibility to 100
        Check instances na_image
                { Calculate space_na_image
                   total_area_image – total_area_na_image   //  total area of images minus
                total area of non accessible images
                        if  (space_na_image is greater than 75)
                                then reduce accessibility by 75
```

```
                    if (space_na_image is greater than 50 but less than 75)
                            then reduce accessibility by 50
                    if (space_na_image is greater than 25 but less than 50)
                            then reduce accessibility by 25}
        {Check usage of frames
                    if (usage of frames)
                            then reduce accessibility by 15}
        Check usage of tables
                    if (non-linear)
                            then reduce accessibility by 10}}
```

5.3 Algorithm for Motor Restricted Users

```
If (user = 'Motor-restricted')
{       set page accessibility to 100
        {Check for a logical tab order through links
                if (no logical tab order)
                        then check for keyboard shortcuts
                                if (no keyboard shortcut)
                                        then reduce accessibility by 60
            {Check scroll bar
                if (scroll bars)
                        then reduce accessibility by 30}}
```

5.4 Algorithm for Dyslexic Users

```
If (user = 'Dyslexic')
{       set page accessibility to 100
        {Check font size   // This would apply to main body text
                if (font size is fixed)
                        then reduce accessibility by 20
                if (font size of body text is less than 10pt)
                        then reduce accessibility by 20
                if (font size of body text is less that 12pt but greater than 10pt)
                        then reduce accessibility by 10
        Check for headings //  Checking for large blocks of text *}
        { Check background colour and font colour//  body text
                if (background colour is white and font colour is black)
                        then reduce accessibility by 20
                if(background colour is black and font colour is white)
                        then reduce accessibility by 20}}
```

6 How the System Works

The system will incorporate a user model and web model with new attributes that are particularly relevant to users with special needs. The information in the user model is used to determine which algorithm is used to rate the accessibility of a web page. The type of algorithm applied will depend upon whether the user is: blind; short-sighted; motor restricted or dyslexic. The user supplies this information. A database acts as the web model, which represents the accessibility of each page and includes attributes

such as the URL and accessibility rating of web pages. Each web page is then analysed according to the constraints incorporated within the algorithm. When a user submits a query, the system refers to the user model to establish which algorithm is assigned. It then individually analyses the first 20 web pages that the search engine returns, and inputs the URL and the rating of accessibility it has given the page into the database. It is hoped that by restricting the rating of web pages to only 20, it is then less likely that accessibility will override content. Following this, the system then ranks the web pages according to the ratings applied and then presents the 20 pages as a list of URLs to the user, according to the ratings that have been assigned. Pages with high ratings are considered more accessible than those with low ratings.

6.1 How the System Works for a Specific User

Using short-sighted users as an example, the algorithm that would be applied will specifically look at font sizes and the contrast between font and background colour. Although the results indicate that users specified advertisement pop-ups as being the third element to affect accessibility, locating the presence of this element has proved to be problematic. This is because it is difficult to determine whether the HTML relating to the pop up explicitly relates to an advert or an instance of another window opening up. A manual check would look at embedded scripts, included scripts, and event handlers for "window.open()" methods, however this is beyond the scope of the system's capabilities. It is hoped that further research will allow for this feature to be incorporated.

Each web page is initially assigned a default rating of 100. The rating will decrease depending upon the extent to which affecting elements are present. For the short-sighted user, the algorithm will firstly check that web pages are not designed with a fixed font. The system considers this to be the most important feature when analyzing web pages for font size because most short-sighted users will need to increase the font. Pages that have a fixed font will then receive a higher penalty. The algorithm is then designed to look at the size of the font used for the body text. Only pages that use a font size below 12pt will be penalised. The next element that is checked will be colour contrast. According to the formula provided by the W3C, the range for colour difference needs to be above 500 if the text is to be readable:

> *Color difference is determined by the following formula: (maximum (Red value 1, Red value 2) - minimum (Red value 1, Red value 2)) + (maximum (Green value 1, Green value 2) - minimum (Green value 1, Green value 2)) + (maximum (Blue value 1, Blue value 2) - minimum (Blue value 1, Blue value 2))* (W3C, 2000)

The algorithm is designed to assign penalty points according to the extent of colour difference. Those pages with very low colour difference will receive a lower acessibility rating than those with a higher difference. These algorithms will be used to analyse and select web pages according to user preferences. An additional study will then evaluate the effectiveness of the presentation order of web pages. The study will also evaluate the results from the exploratory study for consistency. Further research will then be carried out to include all elements and then to develop more complex algorithms based upon the results of this.

7 Conclusion

Where many current methods tend to focus upon the adaptation of content, the proposed approach is to adapt the order of search results according to how accessible each page is to the user. Web pages will be assessed according to individual needs using user-defined ratings of elements that affect accessibility. The more accessible the web page is, the higher up the list of results it will feature. This is a far more simple approach than modifying content; a process that retains a certain amount of fallibility due to its dependence upon correctly coded pages. This proposed scheme would reduce the need to sift through numerous inaccessible pages to locate required content, which can be particularly arduous and frustrating task for an impaired user. This could then improve the effectiveness of the World Wide Web as a pedagogical tool for this particular user group.

Since this research commenced, Google has devised a means of selecting web pages specifically for visually impaired users. Their aim is to reduce the amount of inaccessible pages that visually impaired users are forced to sift through to find necessary information. Named 'Google Accessible Search' (Google, 2007), it searches for pages using its search algorithms and couples this by applying additional factors to rank the pages. This process involves looking at HTML markup. 'Google Accessible Search' assesses the overall simplicity of a page. The extent of imagery is examined as well as establishing how the page is affected by disallowing images. In addition, it will assess whether the page facilitates keyboard navigation (Google FAQ, 2006). The method that is proposed in this paper works on a similar principle, in that pages are ranked according to how accessible they are. Both systems use algorithms to rank pages according to factors of accessibility and whilst the factors are very similar for blind users, the proposed system here will also consider elements that specifically affect short sighted users such as contrast and fixed fonts. The other difference is that the proposed system is not limited to visually impaired and blind users but also considers dyslexic, and motor restricted users.

References

Deets, H.B: Aging and Technology: The Convergence of Two Revolutions. CyberPsychology & Behavior 2(6), 501–503 (1999)

Disability Rights Commission The Web: Access and Inclusion for Disabled People. London (2004)

Henze, N.: Open Adaptive Hypermedia: An Approach to Adaptive Information Presentation on the Web (2001)

Kobsa, A., Stephanidis, C.: Adaptable and Adaptive Information Access for all users, Including Disabled and Elderly People. In: Proceedings of the 2nd Workshop on Adaptive Hypertext and Hypermedia. Hypertext'98. Pittsburgh, USA (June 20-24, 1998)

Good, A.: Enabling Accessibility and Enhancing Web Experience: Ordering Search Results According to User Needs. In: Proceedings of HCI International 2005 (2005)

Good, A.: Addressing User Needs: Adapting Information Access For The Elderly. HCI 2004, HCI and the Older Population Workshop (2004)

Google Accessible Search FAQ. Google labs. (Retrieved January 11, 2007) from http://labs.google.com/accessible/faq.html

Nielson, J.: Beyond Accessibility: Treating Users with Disabilities as People. (Retrieved December 12, 2004) (2001) http://www.useit.com/alertbox/20011111.html
National Statistics Omnibus Survey Internet Access. Individuals and Households (2003)
Preece, J., Rogers, Y., Sharp, H.: Interaction Design. John Wiley & Sons, New York
Shneiderman, B.: Universal Usability. Communications of the ACM. vol. 43(5) (May 2000)
Skinner, B.F.: Teaching Machines. Science. vol. 128 (1958)
W3C/WAI Web Content Accessibility Guidelines 2.0. (Retrieved December 14, 2004) (2004) http://www.w3.org/TR/2004/WD-WCAG2020040311/
W3C/WAI How People With Disabilities Use the Web. (Retrieved November 18, 2004) (2004) http://www.w3.org/WAI/EO/Drafts/PWD-Use-Web/Overview.html
W3C Techniques For Accessibility for Evaluation and Repair Tools. (Retrieved March 28, 2005) (2000) http://www.w3.org/TR/AERT

An Evaluation of Accessibility of Hierarchical Data Structures in Data Broadcasting -Using Tactile Interface for Visually-Impaired People-

Takuya Handa, Tadahiro Sakai, Kinji Matsumura,
Yasuaki Kanatsugu, Nobuyuki Hiruma, and Takayuki Ito

NHK Schience & Technical Research Laboratories,
1-10-11, Kinuta, Setagaya-ku, Tokyo, Japan
{handa.t-es,sakai.t-jk,matsumura.k-ke,kanatsugu.y-ki,
hiruma.n-dy,itou.t-gq}@nhk.or.jp

Abstract. We have been developing a barrier-free information receiving system for the purpose of communicating information in digital broadcasting to visually-impaired people. In the service of data broadcasting in digital broadcasting, many items constitute the menu screen. In this report, presentation methods and access methods using touch in combination with audio for the purpose of effective communication of menu screen structure are briefly explained. Secondly, results of evaluation experiments conducted to obtain a guideline to design hierarchical presentation structures easily accessible by visually impaired people, using the tactile interface in combination with audio presentation were discussed by focusing on the hierarchical structure of menu screens in data broadcasting.

Keywords: accessibility, hierarchical structure, data broadcasting, tactile interface, visually-impaired.

1 Introduction

TV is one of the most basic methods to provide information in our society; therefore it is an important task to develop an information presentation method as well as an access method easily understood by everyone. The authors of this paper have been developing a barrier-free information receiving system for the purpose of communicating information in digital broadcasting to visually-impaired people including those totally blind, with weak vision and blind-deaf. We are also researching optimum information presentation methods as well as access methods based on information reception and recognition [1]. In the service of data broadcasting in digital broadcasting, many items including news and weather reports constitute the menu screen. In addition, since the menu screen has a hierarchical display in many cases, it is not easy for visually impaired people to promptly understand the screen composition and hierarchical structure by audio guidance only to access desired information [2][3]. It is therefore implied that parallel use of the

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 45–54, 2007.
© Springer-Verlag Berlin Heidelberg 2007

audio presentation and tactile presentation is effective for understanding the spatial structure of the screen. [4][5]. In this report, presentation methods and access methods utilizing touch for the purpose of effective communication of menu screen structure and electronic program guide (EPG) in data broadcasting are briefly explained. Secondly, results of evaluation experiments conducted to obtain a guideline to design easily accessible contents by visually impaired people, using tactile interface in combination with audio presentation were discussed by focusing on the hierarchical structure of menu screen in data broadcasting.

2 Hierarchical Structures and Presentation Methods in Data Broadcasting and the Web

Data broadcasting and EPG are operated with GUI that accesses information by selecting a desired item on the screen. In many cases, each item in the menu screen in data broadcasting is indicated hierarchically by category. In the case of the PC and the Web are composed of complicated hierarchy or nesting. For access methods of PCs for visually impaired people, the interface with sequential information using arrow key operation and audio (audio navigation) is the mainstream. However, it is not easy for visually impaired people to understand the screen composition and its hierarchical structure by audio guidance only. Especially for the TVs, it is an important task to develop information presentation and access methods easily understood by everyone. We therefore have been developing a barrier-free information receiving system. In the previous studies on the accessibility of hierarchical structures, it is pointed out that the factors such as the overall site structure and understandability of labels would influence the accessibility [6]. On the other hand, it is indicated that search methods with tactile display are effective because they make search paths fewer and estimating approximate position of target objects easier [4]. As an example of tactile navigation method that use touch in combination with audio presentation, we have developed an array-type tactile interface that incorporates tactile keys having an projectile/retractile mechanism. We call the proposed system Tactile Navigator [1].

3 Communication of Menu Screen Structure in Data Broadcasting by the Tactile Navigation

3.1 Composition and Moving Function of Tactile Navigator

In Tactile Navigator (Figure 1), vertically moving switches (tactile keys) are installed in the arrangement of telephone buttons because it is known that the arrangement is suitable for daily use, reflecting the results of the preliminary experiments. In the preliminary experiments, time required for key search was obtained depending on the arrangement of tactile keys as well as the height of keys. As a result, the search time was significantly shorter in the case of the arrangement of telephone buttons than in the case of one horizontal column such as numerical keys in a full keyboard,

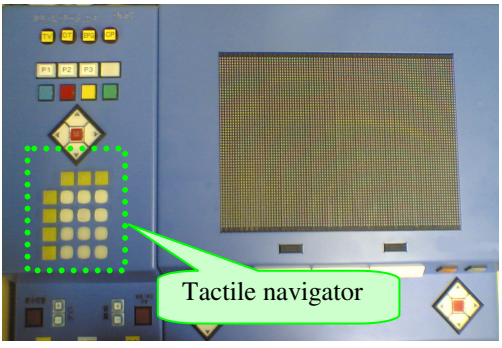


Fig. 1. Tactile Presentation Terminal

and the best height of keys was approximately 3mm.

In this developed system, users are able to know the selectable items in the GUI screen as well as composition of cells in a table such as EPG, by touching the presentation surface of Tactile Navigator and understanding the arrangement, shape and the number of tactile keys. By pressing down a tactile key, the content can be confirmed by audio or Braille, to access information [1].

3.2 Tactile Presentation of Menu Screen Composition in Data Broadcasting

Methods to present items originally in a GUI screen onto a tactile display include a screen presentation method which faithfully presents the shape of screen items as well as a logical presentation method which presents the items by corresponding them to other objects with different abstract shapes and sizes for easy tactile recognition [1]. As shown in Figure 2, Tactile Navigator is a logical presentation method. The figure shows an example of the operation of Tactile Navigator. The raised circular tactile keys S1-S7 represent the number of selectable items, and the raised square tactile keys N1-N3 represent levels in the menu hierarchy. The remaining tactile keys (shown as dotted lines) are at the same height as the background surface. By using differently shaped key for the selectable items and for the hierarchical levels, these different capabilities can be easily differentiated by touch. For example in Figure 2, the current location is in the third depth of the hierarchy, indicating that there are seven selectable items. It is also possible to directly transit among layers with the hierarchical level keys [1].

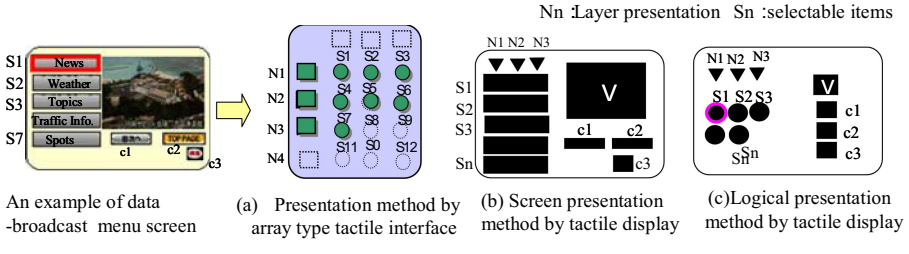


Fig. 2. Presentation method for selectable items and layer structure using tactile unit

4 Evaluation of Accessibility of Hierarchical Structures

4.1 Purpose of Experiments

According to the previous experiments, the tactile navigation proposed by the authors was found to be able to considerably decrease the time to search desired items and to

be effective for accessibility improvement compared to the audio navigation. Furthermore, based on introspective reporting by the subjects, it has been indicated that the tactile navigation is effective not only in terms of operability and accessibility but also for understanding the composition of menu screen or tables in data broadcasting [1]. However, the previous study was conducted on one specific hierarchical structure. Therefore it is necessary to examine more general cases. The purpose of this experiment is to clarify the hierarchical structure easily accessible to menus in data broadcasting. Thus, by focusing on the hierarchical structure of menus in data broadcasting, we conducted evaluation experiments to obtain a guideline for hierarchical presentation easily understood and accessible by visually impaired people.

4.2 Evaluation Experiments in Regard to Presentation of Hierarchical Structures

Experimental Method and Conditions

In the experiments, we evaluated differences in the search time and the number of operations for items given randomly by the audio navigation and the tactile navigation to hierarchical structures with different numbers of layers and branches. The experimental conditions are described in Table1.

Table 1. Experimental conditions

Factors	Levels		
Presentations method	Audio navigation		Tactile navigation
Hierarchical structure	4- branches 4-layers	8- branches 3-layers	16-branches 2-layers

With hierarchical structures and presentation methods as evaluation factors, conditions included three patterns of hierarchical structures and two presentations methods: the audio navigation and the tactile navigation. The simulated structure of data broadcasting menu screens with 4- branches/4-layers used for the experiment is shown in Figure 3. The hierarchical structure was given four layers at maximum, considering the number of layers presented in typical data broadcasting. With the total number of items in the deepest layer (the number of terminal nodes) as 256, three patterns were prepared so that the number of branches equaled 2^n (Table 2). In regard to the 8-branches/3-layers, there are four items only for the third layer, in order to make the number of terminal nodes 256. The variety in the total number of nodes was not adjusted in this experiment.

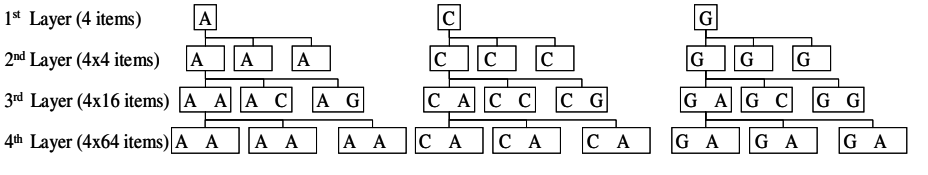


Fig. 3. Example of the hierarchical structure for evaluation experiment

Table 2. Hierarchical structures for experiment

	Number of terminal nodes	Total number of nodes
4- branches/4-layers	256 (4×4×4×4)	340
8- branches/3-layers	256 (8×8×4)	328
16- branches/2-layers	256 (16×16)	272

Items in each hierarchy were composed by the combination of letters of "alphabet" and "hiragana," which were lined up alternatively one letter after another starting from the first letter. As the layer goes deeper, alphabet and hiragana were connected alternatively so that the layer and item can be uniquely identified with a given task. We asked to search an item given randomly without limiting to terminal nodes, and evaluated the understandability and accessibility of hierarchical structures according to the search time and the number of operations. The starting point was the first item in the first layer for the first task of each experimental condition, or the item at the time of completion of a previous task for other cases. The interval between the point when a task was given and the point of task completion was measured as the search time. We did not give subjects instruction such as "as quickly as possible" or "fewest operations as possible" but asked to search freely.

In the experiments, search tasks with 36 items were conducted for all six combinations of factors and levels indicated in Table 1. With this 216 (36 x 6) trials as one cycle, experiments were repeatedly conducted five times, and learning effects to the hierarchical structure were also evaluated. The subjects included two visually impaired people and two non-visually-impaired people (Table 3). The non-visually-impaired people were blindfolded upon experiments.

Table 3. Subject profiles

Subject	Sex	Age	Lost sight era
A	F	20	Sighted
B	F	20	Sighted
C	F	20	10
D	F	50	10

Experimental Results

Regarding the search time and the number of operations, we conducted an analysis of variance (ANOVA) based on a two-way layout using “presentation method” and “hierarchical structure” as the factors. The main effect of each were significant at a 1% level. In Figure 4, 95% confident interval of the sample mean and population mean of the number of operations are shown in terms of hierarchical structures. According to this result, there was a tendency mostly consistent with the theoretical minimum numbers of operations in the tasks. The difference in the number of operations between the audio navigation and the tactile navigation was significant for more branches and fewer layers.

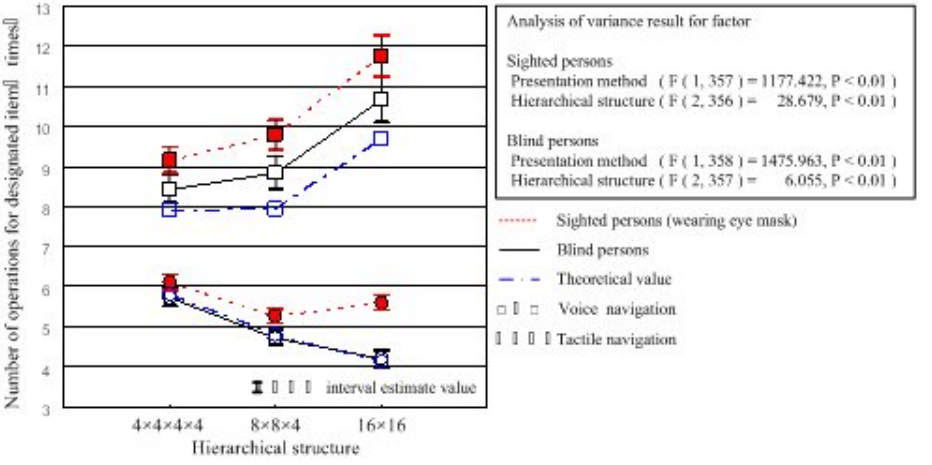


Fig. 4. Number of operations depend on presentation methods and hierarchical structures

Figure 5 shows the results on the search time. In regard to the search time, very similar tendency was observed for both non-visually-impaired people and blind people although the absolute values were different, and tactile navigation with 8-branches/3-layers had the shortest time. The difference in the search time between the audio navigation and the tactile navigation was smaller for more branches and fewer layers. The theoretical value of the shortest operation and search time indicated in the figure assumes that one second is the minimum time necessary for one mechanical operation, representing the product of the minimum time and the number of

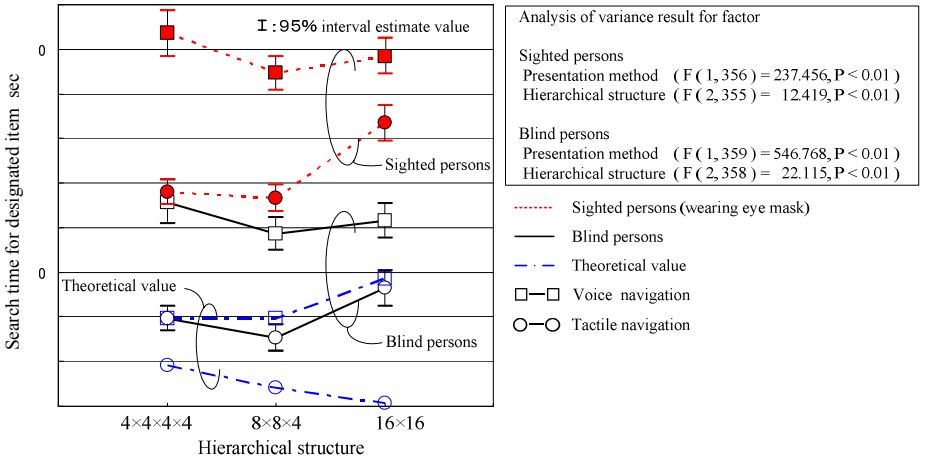


Fig. 5. Search time depend on presentation methods and hierarchical structures

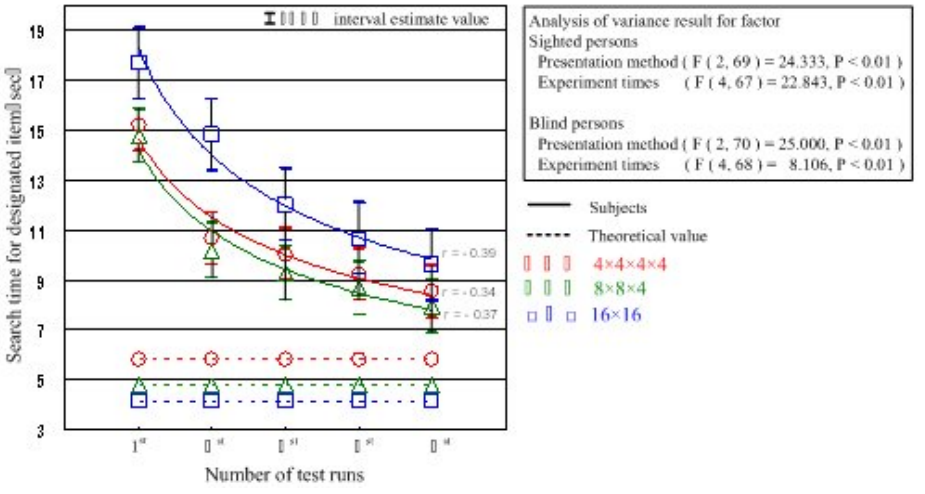


Fig. 6. Search time and learning effect depend on hierarchical structures by using tactile navigation

operations for the shortest route in the given task. Thus, it means the theoretically shortest time in the case that the task is performed with the shortest route without any thinking intervals.

We conducted an ANOVA of the search time treating the hierarchical structure and number of test runs as a two-way layout. The main effects of each were significant at a 1% level. In Figures 6, the horizontal axis shows the number of test runs, and the search time is expressed for each hierarchical structure.

5 Discussion

As a result of evaluation of the influence to the search time and the number of operations by factors of presentation methods and hierarchical structures, both the number of search operations and the length of search time were significantly smaller in the case of the tactile navigation than the audio navigation. This is consistent with the results of evaluation experiments conducted previously [1]. We now discuss understandability and accessibility of the hierarchical structures.

Accessibility of hierarchical structures

According to the results in Figure 5, the search time for 8-branches/3-layers was the shortest for both presentation methods. The 8-branches/3-layers condition is considered as the most accessible structure handled in this experiment. However, according to the prediction from theoretical values in this task, it was expected that 4-branches/4-layers and 16-branches/2-layers would have the shortest search time for the audio navigation and the tactile navigation respectively. Contrary to the prediction, however, 4-brances/4-layers and 16-branches/2-layers had the longest search time for the audio navigation and the tactile navigation, respectively. From

Figure 3, results on the number of operations was consistent with theoretical values in both cases, and it should be noted that increase of the number of operations does not simply cause the lengthening of the search time.

As a factor that caused the longest search time for 4-branches/4-layers in the audio navigation, it is considered that an increase of the number of layers made it difficult to understand the current location and the overall hierarchical structure just by audio presentation. Based on this, it seems that a transition among layers required more time for a decision than a transition among items, which is conducted with the one operation. On the other hand, as a factor that caused the longest search time for 16-branches/2-layers in the tactile navigation, it is considered that the arrangement of items could not be understood and the advantage of direct switching could not be utilized. This seems to be related to the fact that the number of items far exceeds the magic number (7±2) for short-term memory.

As described earlier, despite the fact that the number of operations did not increase even if the number of hierarchies or branches was large (selection was hardly mistaken), the search time for the above two cases was extremely long, which can be considered as the results of longer thinking time. This can be considered as the psychological influence of attempting to memorize the arrangement of items or to avoid unnecessary operations, as well as the influence of the fact that the order of items could be deduced in this task due to its regularity. Further study will be necessary concerning factors depending on users' preferences.

Effects of tactile navigation

In Figure 7, the difference in the search time between the audio navigation and the tactile navigation was described according to the results of Figure 5. This implies that the more different the search time is between the audio navigation and the tactile navigation, the higher the effect of the tactile navigation will be in the hierarchical structure. Based on the results, 4-branches/4-layers was the most effective for the tactile navigation and 16-branches/2-layers was the least. This is considered to imply that the tactile navigation is effective to understand current locations as well as hierarchical structures. Furthermore, in the experiment, cases such that a current location was confirmed several times or the search was continued after returning to the top layer were observed for the audio navigation. On the contrary, according to the introspective reporting by the subjects, the hierarchical structure was easy to understand by utilizing touch. However, it is hard to judge from this experiment whether this effect of the tactile navigation

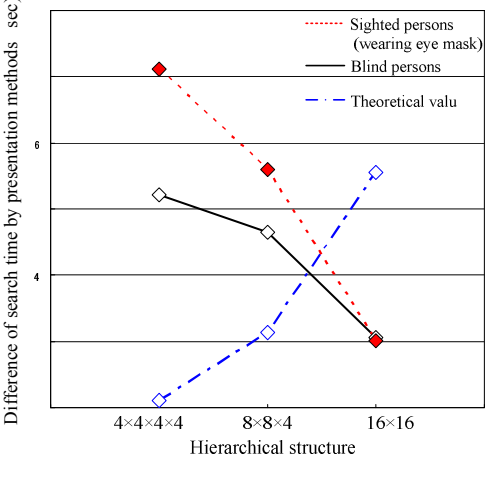


Fig. 7. Difference of search time by presentation methods depend on hierarchical structures

depends on the advantage of direct switching between items/layers or on the tactile presentation itself clarifying the number of items or the current layer position.

Learning effects
The learning curve expressing learning effects can be expressed as the following equation by assuming a negative acceleration type [7].

$$y=A-(A-y_0)\exp(-rx) \quad (1)$$

A: Asymptotic value (limitation of learning)
y_0: Initial value of learning
r: Learning rate (learning speed)

In Figures 6, the learning rate (r) obtained from the learning curve was indicated. In regard to the tactile navigation, significant differences in the hierarchical structure were not observed in terms of learning effects. This can be considered that by combining tactile presentation, a certain learning effect can be expected regardless of the hierarchical structure.

Conditions of favorable hierarchical presentation by tactile interface
It was found in this experiment that approximately 8-branches/3-layers is the favorable condition of hierarchical presentation in tactile navigation, according to the search time for the first trial and subsequent learning effects. It is important to make the hierarchy structure easy to understand and to memorize in order to produce human-friendly contents. These results are considered as important knowledge in a technique to present a menu screen consisting of items as well as the hierarchical structure in data broadcasting.

6 Conclusions

For the barrier-free information receiving system to communicate information provided in digital broadcasting to visually impaired people, we discussed efforts in improving presentation methods from the view point of accessibility. We also evaluated favorable conditions to present hierarchical structure by tactile interface, in regard to items for menu screen or hierarchical structures in data broadcasting. In the experiment, we compared the search time and the number of operations as the evaluation items for the audio navigation and the tactile navigation, mainly from the viewpoint of understanding hierarchical structures. Based on the results, the conditions to present a hierarchy understandable and accessible by visually impaired people were made clearer. However, since the types of hierarchical structures and the number of subjects were limited, further investigation is expected. Since content production and description techniques are so diversified that it is necessary to verify more complicated hierarchical structures and link structures. We also hope to further examine sensory integrative effects as well as effects on short-term memory with touch. We will continue working in order to establish effective presentation methods and access methods for visually impaired people aiming at the information

presentation environment where everyone is able to receive information on broadcasting and communication.

Acknowledgements

This study was partially conducted as part of the "Research and development of multimedia browsing technology for visually impaired people", entrusted by The National Institute of Information and Communications Technology (NICT).

References

1. Sakai, T., Kondoh, S., Matsumura, K., Itoh, T.: Improving Access to Digital Broadcasting for Visually Impaired Users. In: Proceedings of the HCII2005, vol. 7, p. 67 (2005)
2. Ishikawa, J.: Current State and Remaining Issues with Screen Readers for GUI. Transactions of the IPS of Japan 36(12), 1133–1139 (1995)
3. Asakawa, C., Itoh, T.: User Interface for a Non-visual Web Access System. Transactions of the IPS of Japan 40(2), 453–459 (1999)
4. Ebina, T., Igi, S., Miyake, T., Takahashi, H.: GUI Object Search Method by Using a Tactile Display. Transactions of the IEICE of Japan J80-A(11), 2007–2016 (1997)
5. Lee, S.: Access to Computer Systems with Graphical User Interface by Touch. In: Proceedings of the HUMAN FACTORS and ERGONOMICS SOCIETY 39th annual meeting, pp. 742–746 (1995)
6. Aoyagi, S., Hayashi, T., Nakamura, Y., Tondokoro, N.: Accessibility evaluation of website taking into account the use of persons with visual impairments, IEICE of Japan, Technical Report, SP2002-110, WIT2002-50, pp. 9–14 (2002)
7. Umezu, Sagara, Miyagi, Yoda, Psychology Encyclopedia, Heibon co. pp. 85–86 (1981)

Web Accessibility Compliance of Government Web Sites in Korea

Joonho Hyun[1], Doojin Choi[2], and Sukil Kim[3]

[1] Korea Agency for Digital Opportunity and Promotion, 645-11 Deung-chon Dong, Gang-seo Gu, Seoul, Republic of Korea
jhyun22@kado.or.kr
[2] Korea Agency for Digital Opportunity and Promotion, 645-11 Deung-chon Dong, Gang-seo Gu, Seoul, Republic of Korea
djchoi@kado.or.kr
[3] Chungbuk National University, 12 Gashin Dong, Cheongju, Republic of Korea
ksi@cbnu.ac.kr

Abstract. This paper introduces Korean web accessibility activities, such as relational laws, ordinances, policies, guidelines. It also presents analytical result of the investigation on web-contents accessibilities of the 39 Korean government agencies. The result shows that only one agency provides web contents satisfying all the minimum requirements, while 97% of the agencies does not satisfy all the minimum requirements. Unfortunately, 6 agencies do not satisfy any.

1 Introduction

The dramatic development of information technology and internet penetration has brought information society forth. On information society, the source of wealth of the individual, the society and the nation is the ability to create and utilize information and knowledge. Also, internet access and use of it becomes an essential activity. However, the disabled cannot access and use the internet because of his/her physical defects.

Governments around the world have been worried about unequal internet accessibility which deepens the social and economic discrimination around individuals, communities, regions and nations. Global society increasingly depends on the internet to deliver information, enhance communications, improve education and conduct business, so that equal internet access by the individual has become fundamental.

Therefore, Web accessibility has become a main principle in many countries around the world. Developed countries such as U.S, U.K, Australia already have recognized the importance of web accessibility. In order to promote accessibility to the internet and telecommunication devices, they have enacted laws and regulations such as Section 508 of the Rehabilitation Act, DDA (Disability Discrimination Act) and have supported development of assistive technologies [5].

However, Korean government had had a little concern about web accessibility. In 2002, Ministry of Information and Communication (MIC) enacted 'Recommendable Guidelines to improve accessibility for the disabled and the elderly to the IT services

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 55–63, 2007.
© Springer-Verlag Berlin Heidelberg 2007

and products. Recommendable guidelines have not been effective because they were not compulsory, such as obligations like Section 508.

This paper presents policies and standardization of Korean government for web accessibility, as well as a survey of government web sites in Korea. The web sites for the 39 government agencies were examined using KADO-WAH, a web accessibility evaluation tool such as Bobby™, and subsequent manual tests for 4 major checkpoints among 14 Korean web accessibility checkpoints to determine their compliance of web accessibility.

2 Overview of Korean Web Accessibility Legislations and Standard

Korean government has already recognized the problem caused from the digital divide. The former president Kim, Dae-Jung suggested several projects to solve the digital divide in Korea and Asia at ASEM summits in year 2000, APEC summits in 2001 and ASEAN+3 summits in 2001. Furthermore, MIC made a master plan and executed several projects to bridge the digital divide. Since 2002, Korea has made an effort to improve web accessibility [4].

2.1 Web Accessibility Policies in Korea

Korean government has taken active steps toward solving the digital divide for the disabled. Polices for improving IT usage of the disabled are classified into 6 groups: developing assistvie technology for disabled people, providing IT learning opportunities, developing and distributing online contents, enacting laws and regulations for the disabled, constructing community access center, and distributing recycled personal computers.

In order to improve the web accessibility for the disabled, the government has formulated three important policies [1], [4]. Firstly, the government enacted laws and regulations. In 2001, the National Assembly of Korea enacted ‘The Digital Divide Law’ and implemented "Master Plan for Closing The Digital Divide (2001~2005)". In this initiative, the following four action plans were suggested: to provide at least one free internet accessible site in each local administration office (about 3,500 units), to train all the people who wish to learn the advanced as well as basic IT skills, to spread high-speed internet services throughout the nation including the remote areas by 2005, to assist the development of valuable online information for the disabled and the elderly.Second government policy is to support operation of the Information and Telecommunication Accessibility Promotion Standard Forum (http://www.iabf.or.kr) which is founded in 2002. The purpose of the forum is to promote accessibility to telecommunication devices and telecommunications services through the sharing of relevant information among developers and scholars in this field. This forum is doing five major activities to achieve this purpose.

1. Development of technical standards for improving accessibility to telecommunication devices and information service.

2. Sharing information on the trends of accessibility policies, laws and regulations and relevant technologies such as voice recognition and synthesis, digital talking book, various new technologies related to accessibility issues.
3. Development of methods and tools needed to implement accessibility evaluation on telecommunication devices and information services.
4. Cooperation with international organizations which deals with accessibility issues.
5. Promotion of public awareness on accessibility items through holding seminars and workshops, making brochures and operating homepage.

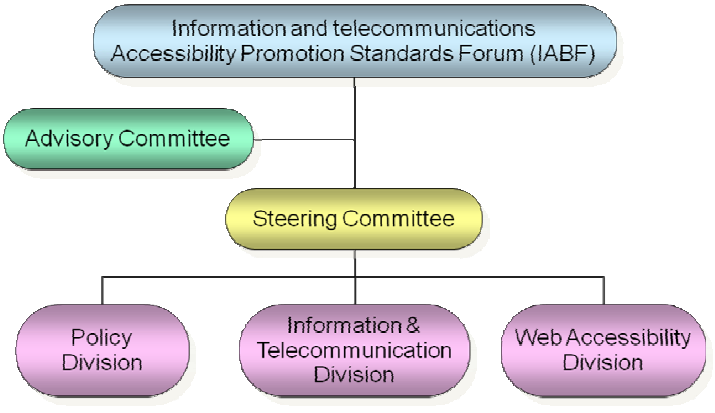


Fig. 1. Organization of Information and telecommunication Accessibility Promotion Standard Forum

Results of activities today since 2002, the states of accessibility technologies and products have been thoroughly reviewed. Also, IABF forum developed web accessibility evaluation tools such as Korean version of 'A-Prompt' in 2003 and "KADO-WAH" in 2005.

Third government policy is to support for the disabled people to buy assistive technologies such as screen reader, Braille display, telephone for hearing-impaired people, etc. Also, government has supported to develop assistive technologies for disabled people.

2.2 Korean Web Contents Accessibility Guidelines

The Korean Web Content Accessibility Guideline 1.0 (KWCAG 1.0) outlines design principles for creating accessible web contents. KWCAG 1.0 had adopted guidelines of Section 508 § 1194.22, a standard for accessible web-based intranet and internet information and applications.

Similar to Section 508 § 1194.22, KWCAG1.0 puts into level 1 guidelines of WCAG1.0, a W3C prepared guidelines

KWCAG 1.0 also had referred to Core checkpoints of W3C WCAG 2.0 as of June 2003 that would be the minimum requirements. KWCAG 1.0 categorizes 14 checkpoints into 4 basic principles, such as perceivable, operable, understandable and

Table 1. 14 Checkpoints of KWCAG 1.0

#	Checkpoint
1	All non-text content that can be expressed in words has a text equivalent of the function or information that the non-text content was intended to convey
2	For any time-based multi- media presentation, synchronize equivalent alternatives with the presentation
3	Web pages shall be designed so that all information conveyed with color is also available without color, for example from context or markup
4	If image maps are used in the content, use client-side image map only or server-side image map with redundant text links
5	Use small number of frames of every web page, and title each frame to facilitate frame identification and navigation
6	Pages shall be designed to avoid causing the screen to flicker
7	All functionality is operable at a minimum through a keyboard or a keyboard interface
8	A method shall be provided that permits users to skip repetitive navigation links
9	When a timed response is required, the user shall be alerted and given sufficient time to indicate more time is required
10	Data table should be designed to provide enough information about corresponding cells
11	Structures and/or formats of documents shall be organized so they are understandable

Table 1. (*Continued*)

#	Checkpoint
12	When electronic forms are designed to be completed on-line, the form shall allow people using assistive technology to access the information, field elements, and functionality required for completion and submission of the form, including all directions and cues
13	Content elements, such as script, plug-ins, applets or other application can be used only when they can provide equivalent information using the current assistive technology
14	A text-only page, with equivalent information or functionality, shall be provided to make a web site comply with the provisions of this part, when compliance cannot be accomplished in any other way. The content of the text-only page shall be updated whenever the primary page change

robust. Perceivable web content implies that all content can be presented in the form that can be perceived by any user [1], [2]. Operable web content ensures that the interface elements in the content are operable by any user. Web content and controls that are easy to understand, they are known to be understandable. Robust web content is to use web technologies that maximize the ability of the content to work with current and future assistive technologies and user agents.

3 Web Accessibility Compliance of Korean Government Web Sites

3.1 Evaluation Methodology

In order to examine web accessibility compliance status of Korean government web sites, 39 agencies were chosen according to Cheong Wa Dae (presidential office) web site's executive branch. This study was conducted in July 2006 in order to evaluate web accessibility compliance of the government web sites.

The KWCAG 1.0 has 14 checkpoints. However, only 4 major checkpoints, that must be the crucial among 14 checkpoints, were considered to evaluate 39 agency web sites. 4 checkpoints are summarized in Table 2. To save the resources, only the main page (home page) of the web sites is evaluated during the evaluation.

Table 2. Minimum requirements used for the web accessibility evaluation

Checkpoint	Minimum Requirement (Pass Criteria)
Checkpoint 1 (Alternative text)	Alternative text of all non-text contents must be provided more than 90% and every alternative text must be a textually equivalent to the function or information.
Checkpoint 5 (Frame title)	The number of frame of a page must be less than 4 and each frame must have adequate name.
Checkpoint 7 (Keyboard)	All functionality must be operable at a minimum through keyboard.
Checkpoint 11 (Logical Structure)	Structures and/or formats of documents must be organized to be understandable.

Two processes have been accomplished: one is an automatic test by using an accessibility evaluation tool, and the other is a manual test to inspect undetected elements by the automatic test. In the automatic test, KADO-WAH, which was developed by Korean Agency for Digital Opportunity and promotion (KADO), was used to evaluate the existence of alternative text for non-text content, and to evaluate the existence of the title of the frame. Manual test was conducted to look at the source codes.

3.2 Results

Table 3 shows the results of our study. 39 government agencies in Korea were violating KWCAG checkpoints. Only Ministry of Finance and Economy were conforming to 4 checkpoints of KWCAG 1.0 shown in Table 2. 38 agencies did violate at least one checkpoint of Table 2. Only 34% of the agencies satisfies the checkpoint 1 (Alternative text), 31% satisfied the checkpoint 2 (Frame), 26% satisfied the checkpoint 3 (Keyboard), and 64% satisfied the checkpoint 4 (Logical structure). According to the study, 6 government agencies did meet none of the minimum requirement. They are Ministry of Agriculture and Forestry, Ministry of Commerce, Industry and Energy, National Policy Agency, Rural Development Administration, Small and Medium Business Administration, and Industrial Property Office.

Table 3. Summary of web accessibility evaluation result (2006. 7)

Agency name	Chckpnt 1 (Alt-text)	Chckpnt 5 (Frame)	Chckpnt 7 (Keybrd)	Chckpnt 11 (Logical)
Ministry of Finance and Economy				
Ministry of Education and Human Resource Development				
Ministry of Science and Technology				
Ministry of Unification				
Ministry of Foreign Affairs and Trade				
Ministry of Justice				
Ministry of National Defense				
Ministry of Government Administration and Home Affairs				
Ministry of Culture and Tourism				
Ministry of Agriculture and Forestry				
Ministry of Commerce, Industry and Energy				
Ministry of Information and Communication				
Ministry of Health and Welfare				
Ministry of Environment				
Ministry of Labor				
Ministry of Gender Equality				
Ministry of Construction and Transportation				
Ministry of Marine Affairs and Fisheries				
Ministry of Planning and Budget				
Government Legislation Agency				
Government Information Agency				
Patriots and Veterans Administration Agency				
National Tax Service				
Customs Service				
Supply Administration				
National Statistical Office				
Meteorological Administration				
Supreme Public Prosecutor's Office				
Military Manpower Administration				
Defense Acquisition Program Administration				
National Policy Agency				
National Emergency Management Agency				
Culture Properties Administration				
Rural Development Administration				
Forest Service				
Small and Medium Business Administration				
Industrial Property Office				
Food and Drug Administration				
Fair Trade Commission				

4 Conclusions

This study evaluated the web sites of government agencies in Korea. The result shows that 97% of the agencies do not conform to 4 major checkpoints of KWCAG 1.0. Only one agency was satisfying all the major checkpoints and only 34% of the agencies meet the checkpoint 1 (Alternative text).

Recently in Korea, associations of the disabled lift up their voice to enact a compulsory regulation for web content providers to make their web pages accessible. As a result, the government has recognized the importance of the web accessibility. Korean National Assembly has been enacting Disability Discrimination Act. Such a movement and efforts would let KWCAG 1.0 be a mandatory regulation in government sectors near soon.

Korean government also closely observes the development of WCAG 2.0. We believe the government will start on revision of KWCAG 2.0 soon after WCAG 2.0 becomes an international guideline.

References

1. Joonho, H., Doojin, C., Sukil, K.: An active step toward a web content accessible society. In: International Cross-Disciplinary Workshop on Web Accessibility(W4A Workshop) at WWW2005 Conference, Chiba, Japan (2005)
2. IABF, Korean Web Contents Accessibility Guidelines 1.0. (December 2004) (2007. 2.14) from http://www.iabf.or.kr/
3. KADO, Web accessibility Compliance Report, Korea Agency for Digital Opportunity and promotion (2006)
4. Seongil, L., Doojin, C., Joonho, H.: Public Awareness on Web Accessibility in Korea. In: The 7th International Conference of Human Services Information Technology Applications(HUSITA7), Hong Kong (2004)
5. Section 508 of Rehabilitation Act (2007.2.14) from http://www.section508.gov
6. W3C, Web Content Accessibility Guidelines 1.0. (May 1999) (1999) from http://www.w3c.org/TR/WCAG10/

Appendix 1: Evaluation Web Sites

Agency name	Websites
Ministry of Finance and Economy	http://www.mofe.go.kr/
Ministry of Education and Human Resource Development	http://www.moe.go.kr/
Ministry of Science and Technology	http://www.most.go.kr/
Ministry of Unification	http://www.unikorea.go.kr/index.jsp
Ministry of Foreign Affairs and Trade	http://www.mofat.go.kr/mofat/index.jsp
Ministry of Justice	http://www.moj.go.kr/
Ministry of National Defense	http://www.mnd.go.kr/
Ministry of Government Administration and Home Affairs	http://www.mogaha.go.kr/gpms/index.jsp
Ministry of Culture and Tourism	http://www.mct.go.kr/index.jsp
Ministry of Agriculture and Forestry	http://www.maf.go.kr/index.jsp
Ministry of Commerce, Industry and Energy	http://www.mocie.go.kr/index2.html
Ministry of Information and Communication	http://www.mic.go.kr/
Ministry of Health and Welfare	http://www.mohw.go.kr/

Ministry of Environment	http://www.me.go.kr/
Ministry of Labor	http://www.molab.go.kr/
Ministry of Gender Equality	http://www.mogef.go.kr/
Ministry of Construction and Transportation	http://www.moct.go.kr/
Ministry of Marine Affairs and Fisheries	http://www.momaf.go.kr/
Ministry of Planning and Budget	http://www.mpb.go.kr/servlet/introScreenServlet
Government Legislation Agency	http://www.moleg.go.kr/
Government Information Agency	http://www.allim.go.kr/
Patriots and Veterans Administration Agency	http://www.mpva.go.kr/
National Tax Service	http://www.nta.go.kr/
Customs Service	http://www.customs.go.kr/
Supply Administration	http://www.sarok.go.kr/
National Statistical Office	http://www.nso.go.kr/
Meteorological Administration	http://www.kma.go.kr/intro.html
Supreme Public Prosecutor's Office	http://www.sppo.go.kr/kor/index.jsp
Military Manpower Administration	http://www.mma.go.kr/
Defense Acquisition Program Administration	http://www.dapa.go.kr/
National Policy Agency	http://www.police.go.kr/main/index.do
National Emergency Management Agency	http://www.nema.go.kr/
Culture Properties Administration	http://www.cha.go.kr/korea/newIndex.action
Rural Development Administration	http://www.rda.go.kr/
Forest Service	http://www.foa.go.kr/
Small and Medium Business Administration	http://www.smba.go.kr/main/index.jsp
Industrial Property Office	http://www.kipo.go.kr/kpo2/user.tdf?a=user.main.MainApp
Food and Drug Administration	http://www.kfda.go.kr/
Fair Trade Commission	http://www.ftc.go.kr/

Usability Evaluation of Korean e-Government Portal

Seongil Lee[1] and Joo Eun Cho[2]

[1] Sungkyunkwan University, Department of Systems Management Engineering,
Suwon, 440-749, Korea
silee@skku.edu
[2] Kyungpook National University, Department of Sociology,
Daegu, 702-701, Korea
june@knu.ac.kr

Abstract. This paper reports the results of an evaluative study of the Korean e-Government portal from the usability perspectives, using a common set of performance metrics from user participation. For fourteen civil affairs service sites, 51 participants performed usability tests and post-test interviews. The results showed that even though the e-Government services are fairly attainable, it takes too many steps and long time for users to complete the given tasks. The overall success rate to complete the tasks on e-Government was 0.69 and it took more than 4 minutes and 12 page navigations on average to complete simple transactions. A strong digital divide could be observed in the use of e-Government services in that the users with different vocations and ages resulted in difference in perceived usability and actual performance. Results suggest that not only are there wide variations in the usability for the services provided, but that significant work still needs to be undertaken in order to make the services of the e-Government more usable, particularly for the older users.

1 Introduction

The rapidly changing information and communication technology forces government agencies to deal with state-of-art technology in order to enhance the agencies' services to the public. The technology presents new opportunities for government to address problems in many conventional services. E-Government embraces the use of modern information and communications technologies to provide information about public services and enable the citizen to conduct transaction electronically with those services. As the use of ICTs becomes more widespread, there is a need to raise the levels of access and participation to ensure that e-Government delivers for all citizens. Provided the service is designed properly, the online delivery may be more user-friendly than that delivered by traditional means [6]. Success of the e-Government portal will largely depend on their accessibility, usability and quality. It is mainly to transform government to be more citizen-oriented [1]. Government exists for the benefit of the citizen and there is a natural expectation that the government recognizes this relationship in the way in which it interacts with its citizens.

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 64–72, 2007.
© Springer-Verlag Berlin Heidelberg 2007

1.1 Objectives of the Study

In recent years there has been worldwide increasing pressure for better government, not just in more efficient and effective way but in better customer services. The need to provide a better quality services, focus on the citizen, embrace more efficient and effective work practices and improve its administrative processes motivated governments throughout the world to develop e-Government. At the same time, e-Government presents new opportunities for improved efficiency and effectiveness to deliver government services in an increasingly complex and changing environment. However, the effort to provide a better e-Government should be felt by its citizens to be successful. The purpose of this paper is to describe the issues related to the usability of Korean government web portal as benchmark measures of good e-Government practices, associated with the digital divide issue. Our goal is to evaluate the usability of the e-Government portal as the main theme of the centrality of the citizen in the provision of the services.

1.2 Related Research

UNDP defined the e-Government as "government's use of ICTs to work more effectively, share information and deliver better services to the public" [10]. A Gartner Group report described e-Government as "the continuous optimization of service delivery, constituency participation, and governance by transforming internal and external relationships through technology, the internet, and new media.". The e-Government Act of 2002 defined the e-Government as "the use by the government of web-based Internet applications and other information technologies, to (A) enhance the access to and delivery of Government information and services to the public, other agencies, and other Government entities; or (B) bring about improvements in Government operations that may include effectiveness, efficiency, service quality, or transformation" (116 Sta. 2899, at 2902). While detailing the definition of e-Government in its contextual aspects, Seifert and Relyea [9] argued that governments are obligated to focus on equality and accessibility to public services and information while representing the interests of the citizenry as a whole. Pascual [8] insisted that e-Government ultimately aims to enhance access to and delivery of government services to benefit citizens. As we can see from all the definitions made from many different studies, the concept of government services to its citizen is in the core of the e-Government.

As in the previous studies on the development and evaluation of web sites [5, 7], the question that drives this study is to determine and increase the understanding of how usable an e-Government portal is for its citizen. It is argued that those charged with implementing e-Government need to give careful consideration to the perceptions and expectations of its users to provide more user-friendly services. Lack of technological competence on the part of users, such as a complex interface putting off citizens from using the service, is regarded as one of user-culture barriers to e-Government service development [11]. Vassilakis et. al. [11] also argued that the two most important factors to lead the e-Government to success are legislative issues and user-related issues, where the latter should be addressed at the service design stage to develop more user-friendly interfaces.

It is important to recognize that the citizen need not understand the complex structures of government in order to access the exact information or to get any government service as required in the traditional off-line services. Instead, e-Government will enable the citizen to access the exact information or government service required through a single electronic gateway. This will reduce the time and effort required for citizens to comply with government rules and regulations [4]. The key recognition within the government sector is that e-Government must be seen in the context of organizing around the citizen. In the center of these arguments are the accessibility and usability of the e-Government portal.

However, accessibility and usability are often disregarded in developing and organizing e-Government sites, and governments seldom conduct usability evaluation for the citizens, who should be the major customers using the system, ignoring the needs of real users. Even though governments recognize the key aspect of the online delivery of services to citizens is the interaction between the citizen and government, governments seldom collect feedback from citizens regarding its access and use. In other words, e-Government often becomes just a one-way delivery of services. In that sense, it is recommended that government web sites should show citizens that their involvement matters by developing feedback mechanisms and encouraging their use [4].

2 Usability Evaluation

2.1 Korean e-Government Portal

The web sites that make up the Korean e-Government were mainly G4C sites for citizens to access and use for civil affairs services. Among the sites analyzed are those of major ministries' offices and agencies serving crucial functions of government to citizens, such as health, human services, taxation, education and training for jobs, and regulation. These civil affairs service sites are different from the respective government agency web sites in that actual transactions for civil services are available while the government web sites are only for providing information on the administrative agenda and public relations of the agencies.

Weather report services were available both in wired and wireless internet services through mobile networks. This evaluation included a wireless internet service through mobile phones. It was one of a few mobile web sites available among the e-Government services.

2.2 Usability Evaluation

Tasks. A total of 14 tasks for civil affairs services from the e-Government was evaluated. Each task consisted of a specific service provided by a certain ministry of the government. Participants were asked to complete the task unless they were forced to stop after a certain limit of time. The selected tasks were as follows, with the responsible agency of the services in parentheses:

1. Registration and logging-in to the e-Government portal (Ministry of government administration and home affairs)
2. Reporting a lost social security card (Ministry of government administration and home affairs)

3. Application for a copy of one's family register (Ministry of government administration and home affairs)
4. Registration and upload of one's resume to a Work-net (Ministry of labor)
5. Inquiry of a building registration information to buy a house (Ministry of construction and transportation)
6. Inquiry of a land registration information to build a plant (Ministry of construction and transportation)
7. Inquiry of tax arrears (National Tax Service)
8. Registration of one's mobile phone number for cash receipt (National Tax Service)
9. Inquiry of 4 major welfare-related insurance services (Ministry of health and welfare)
10. Inquiry of job training in job training info net (Ministry of labor)
11. Request for an open public information (Ministry of government administration and home affairs)
12. Report to the Cyber Police of a crime (National Police Agency)
13. Weather report
14. Plug-in installation for e-Government program

Subjects. A total of 51 users participated in the evaluation. The participants were selected based upon their age and vocation. All the participants were the residents of Seoul and Daegu metropolitan areas in Korea. About 39% of the participants were college students (n=20), 33% were blue collar workers (n=17), and the rest of them were white collar workers (n=14). The participants were categorized into four age groups ranging from their 20s (n=16), 30s (n=17), 40s (n=10), to 50s (n=8). There were only 8 participants in the 50s since it was hard to recruit the older users who had had previous internet experiences. Twenty seven of the participants were male and twenty four were female.

Procedure. A repeated measure experimental design was used in which each participant performed the tasks for all 14 sites. The order of tasks was counter-balanced among the participants. Participants were given a task at a time with a simple scenario. The tests were conducted at university laboratories in Suwon and Daegu. Two video cameras recorded the participant's face and computer screen to monitor the participant's behavior. Participants were encouraged to speak aloud as much as possible while completing the tasks, but were not forced. Participants' mouse navigations and clicks on the monitor were captured. The time for task completion was also measured as well as the time spent in errors.

After completing the tests, a post-test interview was held for each participant for additional information. Participants were interviewed for their awareness of the e-Government services and previous experiences. Also obtained from the interview were subjective ratings on each participant's satisfaction with the e-Government services. A 5-point scale was used for the rating in each service to examine 1) how usable the participant perceived the service to be, 2) whether the participant would use the e-Government service again in the future, and 3) whether the participant would recommend others to use the e-Government services. Participants could also make specific comments for their opinion on design and usability.

A notable characteristic of the Korean e-Government services is that some of the civil affairs services require the use of electronic certificate of authentication to access the services in addition to initial log-in process. Tasks 2, 3, and 7 require participants to have obtained an electronic certificate of authentication from financial institutes such as banks or from government-certified attestation companies to access and use the e-Government services. It was rather a complicated process for some participants and was not considered in this usability test since the electronic certificate of authentication is beyond the use of e-Government.

3 Evaluation Results

The data for our analysis consist of an assessment of 14 Korean e-Government web sites by 51 subjects. The analysis on the usability evaluation test data was supplemented by subjective rating data from post-test interviews.

3.1 Results of the Usability Tests

The overall success rate for the selected tasks was 0.69 and overall mean time to complete the given tasks was 269.9 seconds. The mean error time, the time for the participants to spend performing unrelated tasks or navigating wrong pages during a given task, was assessed to be as long as 93.7 seconds. The services required 12.1 web page navigations on average to complete the tasks, and 2.5 navigations on average out of the 12.1 navigations were spent for erratic navigations or unrelated manipulations per task. These results suggest that even though the e-Government services are fairly achievable, it takes too many steps and long time for most users to complete the tasks. It took more than 4 minutes on average to complete simple transactions while most users wasted about a minute and a half in error.

Mean success rates and task completion times significantly differed among the fourteen selected tasks ($p < 0.05$).

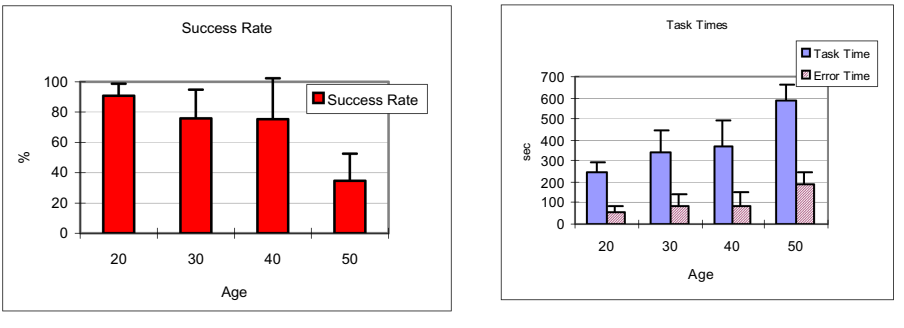


Fig. 1. Mean success rates and task completion times of the participants in four age groups. Participants of the 50 year-old age group showed significantly worse performances than the other groups' participants in all measurements.

A very significant age effect could be observed ($p < 0.01$). Figure 1 shows the mean success rates and task completion times for overall tasks. The participants in their 50s could only finish less than 40% of the given tasks, while the subjects in their 20s could finish approximately 90% of the tasks. The mean task completion time of the subjects in their 20s was observed to be approximately 240 seconds while the participants older than 50 years of age took almost 10 minutes. The mean error time of the participants of 20 year-old age group was shorter than that of the 50 year-old age group by almost 2 minutes. Especially, the tasks of reporting a lost social security card and tax arrears could not be completed at all for the participants in the 50 year-old age group. Both tasks required the use of an electronic certificate of authentication to access which seemed to be hard for certain users both in concept and in acquisition process. No gender effect could be found.

3.2 Post-Test Interview Results

Awareness. Overall awareness of the e-Government services for the given tasks was so low that only 44.1% of the participants were aware of it and only 9% of the participants had previous experiences. Participants in the white collar worker group and college students showed more awareness than the blue collar workers. No difference, however, among the age groups was found.

Perceived Usability. Overall satisfaction rating on perceived usability of the e-Government services was 3.32 out of 5 points. Usability for the tax reporting service for the tax arrears was significantly worse than other services ($p<0.05$). No difference, however, among the age groups nor among the vocation groups could be found. The participants of the 50 year-old age group rated the usability the highest even though they scored the lowest in the actual success rate and task completion time in the usability tests. The participants of the 30 year-old age group rated the perceived usability the lowest at 3.24 out of 5 points. It suggests that the younger users with more experiences in internet use were less satisfied with the usability of the current e-Government system and its interface. Younger participants particularly commented on the difficult terminologies used in the interface while older participants did not like the fonts and size of the letters and menu structures used in the interface.

Intention for Future Use. Overall rating on the intention for future use of the e-Government services was 3.71 out of 5 points. There was not any specific service that the participants rated significantly higher or lower than the others. The participants in the 40 year-old age group rated the highest expressing the intention for future use ($p < 0.05$), followed by the participants in their 50s. The participants of 20 year-old group rated the lowest. The white collar worker group showed significantly more interest in continued future use than the students and blue collar worker groups ($p < 0.05$).

Recommendation to Others. Overall rating on their willingness to recommend the e-Government services to others was 3.79 out of 5 points. The participants of the 40 and 50 year-old groups were significantly more willing to recommend the e-Government services to others than the participants of 20 and 30 year-old groups ($p < 0.05$).

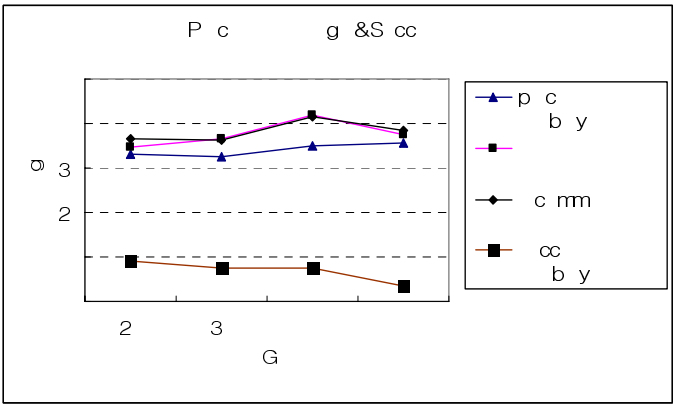


Fig. 2. Mean perceived ratings and success rate in the usability test for the participants' age groups. The aged participants generally showed higher perceived usability, recommendation to others, and intention for future use, while their actual performances in the usability tests were worse than the younger participants.

4 Conclusion

We tried to include all the demographics of the overall population of citizens in this study. Even though the sample size was relatively small, the data reflect a broad diversity of citizens in age, gender, and vocational background. Several issues have been observed that tend to complicate the usability of Korean e-Government.

A strong digital divide could be observed in the actual use of e-Government in that the users with different vocations and ages resulted in difference in perceived usability and actual performance. The difference was so apparent that older users in their 50s had hard time to use the e-Government services even though they seem to be the users who mostly want to use them. The significantly longer task completion times and lower success rates for the older participants clearly showed the usability of the interfaces would act as a clear barrier to the promotion of e-Government services to these users.

As major off-line users of the government civil affairs services, the participants in the 40 and 50 year-old age groups were more willing to use the online version of the government services than the younger participants. It is very typical that the students and participants in their 20s, who would have relatively less experiences in the civil affairs services, were less satisfied and had little intention for future use. They also rated the usability of the e-Government services rather harshly, probably because they have had more experiences with other well-designed sites than the older participants have. Users in their 40s and 50s were less concerned about usability and rated higher in perceived usability, and also rated higher in the intention for continued future use. They might have put blame not on the design of interface but rather on their lack of computer skills. Carter and Belanger [2] found that the perceived ease of use was not a factor for increased use intentions of e-Government services for college students in their study. They also suggested that the experienced internet users such as college

students would have perceived the ease of use differently from the overall population of citizens such as aged users. Findings in our study confirmed, as it can be reasoned, the different perception on the ease of use for the older users. In our study, the perceived usability did not affect both the younger participants' and older participants' intention for continued future use of e-Government services, for different reasons. It was noted that unlike their findings, experienced internet users who may have higher levels of perceived compatibility, the younger participants in this case, were not associated with increased intentions to continuously use the e-Government services in the future simply because they did not understand the immediate needs. With different reasons from Carter and Belanger's study [2], however, lack of perceived ease of use was not a significant deterrent of intention for future use of e-Government for the older users in their 50s, who might have apprehension provoked by complexity in the interface. They did have needs to use the civil affairs services and seemed to understand the real benefits from the online services.

Results also suggested that the usability among the different sites was not evenly distributed. This is probably due to the fact the e-Government portal was simply acting as a mere gate to all the ministries' civil affairs service sites, not being a well-structured portal that should be organized and pre-planned as a whole. It was also not very desirable for users to spend more than 5 minutes on a web site to complete simple transaction tasks, particularly for the aged users who are more willing to use the services on a continued basis. This is partly because many services required the users to input their personal information redundantly, clearly showing the e-Government was not a holistic portal that needs to automatically pass the log-in information to all the services. A person's identity details need only be supplied to government once [4]. It should be noted that one of key goal of e-Government is to reduce the burden of authentication and authorization on the citizen while maintaining a certain level of security. Another factor for the low usability was the use of electronic authentication. Many participants, especially the older participants, were uncomfortable with the electronic authentication, having hard time to understand the concept, to obtain one, to memorize the password, and to repeatedly use it for different services.

It should also be noted that the e-Government portal was designed by private Web design companies under contract with individual government agencies. Without specific government guidelines that dictate the accessibility and usability concern, the design tends to focus more on aesthetics and efficiency by young designers of the private design companies. The inconsistent distribution of usability among the services that we observed during the tests also can be attributed to the lack of consistent guidelines that run through all the government ministries and agencies.

Even though Korean government invested a lot to creating and providing e-Government services to its citizens, awareness and experience of the e-Government services were found to be in a very low level. Even though the Korean e-Government web portal was ranked the highest in the world in the 2006 Brown University's Global E-Government 2006 evaluation [12], lots of Korean users are expressing disappointment in its design and services. It can be understood that it might be very difficult to satisfy the citizens in Korea since they are believed to be one of the most advanced and savvy internet users in the world. It seems like the Brown University evaluation [12] just looked at the number of features and transactions the site

provided online without looking at the actual accessibility and usability. Such an assessment is not desirable since it can misguide the government to be satisfied without giving much effort to put the citizen to the center. While such facts as how many online transactions are provided or how much percentage of its database or video clips are accessible are one thing, the facts on how much effort should be given by citizens in time and steps to actually make a transaction online are another. We regard the latter more important in evaluating the usefulness and value of e-Government.

References

1. Abanumy, A., Al-Badi, A., Mayhew, P.: e-Government Website Accessibility. Depth Evaluation of Saudi Arabia and Oman. Electronic J. e-Government 3(3), 99–106 (2005)
2. Carter, L., Belanger, F.: The Influence of Perceived Characteristics of Innovating on e-Government Adoption. Electronic J. e-Government 2(1), 11–20 (2004)
3. Gartner Group: E-Government strategy: Cubing the circle. Research Notes, Strategic Planning Assumption (April 2000)
4. Information Society Commission: eGovernment: More Than an Automation of Government Services (October 2003)
5. Jarvenpaa, S.L., Todd, P.A.: Consumer Relations to Electronic Shopping on the World Wide Web. International J. Electronic Commerce. pp. 1(2) (1996)
6. Nielsen, J.: Usability Engineering. Academic Press, San Diego (1993)
7. Palmer, J.W., Griffith, D.A.: An Emerging Model of Website Design for Marketing, Communications of the ACM. Communications of the ACM 41(3), 44–51 (1998)
8. Pascual, P.J.: e-Government, e-Primer for the Information Economy, Society and Polity Series. UNDP-APDIP, e-ASEAN Task Force (2003)
9. Seifert, J., Relyea, H.: Considering E-Government from the U.S. Federal Perspective: An Evolving Concept, A developing Practice. J. E-Government. 1(1), 7–14 (2004)
10. UNDP: UNDP Anti-Corruption Practice Note (March 2004) From http://www.undp.org/policy/docs/practicenotes/
11. Vassilakis, C., Lepouras, G., Fraser, J., Haston, S., Georgiadis, P.: Barriers to Electronic Service Development. E-Service Journal, pp. 41–63 (2005)
12. West, D.M.: Global E-Government 2006 Brown University (2006)

Accessibility Evaluation of Korean e-Government

Seongil Lee, Byung Geun Kim, and Jong Gon Kim

Sungkyunkwan University, Department of Industrial Engineering,
Suwon, 440-746, Korea
silee@skku.edu, billkim@limeorage.com, gon0325@gmail.com

Abstract. This paper aims to discover to what extent web accessibility is considered by the e-Government's of Korea. Accessibility of the e-Government of Korea was evaluated using Web diagnostic engines that check the compliance level of Web sites to the W3C guidelines. The accessibility of the e-Government was at a lower level, complying to only 10% of the priority 1 checkpoints. This means that lack of coherent government policy for web accessibility causes lack of providing standard e-Government services for the part of the population with special needs.

1 Introduction

As the technologies of the Information Society present new opportunities to address traditional problems in many conventional services, governments throughout the world have recognized the need to provide a better quality services, focus on the citizen, embrace more efficient and effective work practices and improve its administrative processes by developing e-Government. E-Government embraces the use of modern information and communications technologies to provide information about public services and enable the citizen to conduct transaction electronically for those services. However, there is a real risk that some people will be left behind. Ensuring access for people with disabilities is a key concern of government. There is a clear need to raise the levels of access and participation to ensure that e-Government delivers for all citizens, especially for citizens with disabilities with accessibility provision.

Even though the Korean e-Government web portal was ranked the highest in the world in the Brown University's Global E-Government 2006 evaluation, its W3C disability accessibility profile compliance was only at 15% (Table 1) [6]. Korea was the only country in the top 10 e-Governments of the report that complies with less than 20% of W3C accessibility guidelines, while the average percentage of the other top 10 countries' compliance level reaches almost to 60% in 2006, and 49% in 2005. Korea's accessibility provision in the e-Government was far less than that of the competing Asian countries such as Taiwan, Singapore, Japan, and Hong Kong. The report clearly showed that Korean government did not give much effort to bring citizens with disabilities to its e-Government services, without putting the accessibility and usability issues to the center of all.

This paper reports the results of an evaluative study on the Korean e-Government portal from accessibility perspectives.

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 73–78, 2007.
© Springer-Verlag Berlin Heidelberg 2007

Table 1. Accessibility evaluation of the top 10 e-Government countries (Brown University's Global E-Government 2006 evaluation) [6]

Country	Accessibility 2005 (%)	Accessibility 2006 (%)	Overall Points (/100)
Korea, Republic of	0	15	60.3
Taiwan	92	73	49.8
Singapore	13	20	47.5
United States	44	54	47.4
Canada	70	77	43.5
UK	68	75	42.6
Ireland	89	94	41.9
Germany	50	75	41.5
Japan	35	75	41.5
Spain	30	43	40.6
Average of Top 10	**49.1**	**60.1**	**45.7**

1.1 Objectives of the Study

The purpose of this paper is to describe the issues related to the accessibility of Korean government web portal as benchmark measures of good e-Government practices, associated with the digital divide issue. Our goal is to evaluate the accessibility of the Korean e-Government portal as the main theme of the centrality of the citizen in the provision of the services. It is our intention that the government can improve the accessibility of its services based upon the evaluation.

1.2 Related Studies

Government services are expected to be socially inclusive [3]. This means that all citizens who want to must have ready access to these services. The e-Government Act of 2002 defined the e-Government as 'the use by the government of web-based internet applications and other information technologies, to (A) enhance the access to and delivery of government information and services to the public, other agencies, and other government entities; or (B) bring about improvements in Government operations that may include effectiveness, efficiency, service quality, or transformation.' While detailing the definition of e-Government in its contextual aspects, Seifert and Relyea [4] argued that governments are obligated to focus on equality and accessibility to public services and information while representing the interests of the citizenry as a whole. As we can see from all these definitions made, concept of government services to its citizen is in the core of e-Government. Vassilakis et. Al [5] mentioned accessibility as a component of user-culture barriers, being an obstacle to the use of electronic services of e-Government. Failure to make adequate provision for users with disabilities results in their exclusion from service use.

Accessibility is also an essential core of the services to its citizen. Accessibility is regarded a crucial requirement for successful endeavor for e-Government along with its availability [1]. They argued that the goal of web accessibility is to allow universal access to information on the web, by all people but especially by people with any impairment, no matter what its severity.

It is important to recognize that the citizen need not understand the complex structures of government in order to access the exact information or to get any government service as required in the traditional off-line services. The key recognition with in the government sector is that e-Government must be seen in the context of organizing around the citizen. In the center of these arguments are the accessibility and usability of the e-Government portal. However, accessibility is often disregarded in developing and organizing e-Government sites. Even though government recognizes the key aspect of the online delivery of services to citizens is the interaction between the citizen and government, government seldom collect feedback from citizens regarding its access and use. In other words, e-Government just becomes a one-way delivery of services. Government needs to recognize the accessibility to web an important aspect of quality of life for citizen with disabilities, and to realize the necessity to deliver it for them.

2 Korean e-Government Portal

Korean e-Government portal makes up several pages of ministry services. The site was www.egov.go.kr and consists in Korean. Among the sites analyzed are those major ministries' offices and agencies serving crucial civil affairs functions of government to citizens, such as health, human services, taxation, education and training for jobs, and regulations. Those were mainly G4C sites for citizens to access and use for their civil affairs services.

A total of 7 civil affairs service sites were evaluated. Each site consists of a specific service provided by a certain ministry of the government. The selected sites were as follows with the respective government ministry or agency which is in charge of the services in parentheses:

1. Four major welfare-related insurance services (Ministry of Health and Welfare)
2. Cyber Police Agency
3. e-government portal
4. Job training network (Ministry of Labor)
5. Home tax services (National Tax Service)
6. Inquiry of information openness (Ministry of Government administration and Home Affairs)
7. Tax arrears services (National Tax Service)

We used the software called "KADO-WAH" that was developed by Korea Agency for Digital Opportunity and Promotion to check standard accessibility features of Korean web sites. The program mainly checks if a web site complies to the Korean Web Content Accessibility Guideline 1.0 (KWCAG 1.0). The Korean Web Content Accessibility Guideline 1.0 (KWCAG 1.0) outlines design principles for creating accessible web contents. KWCAG 1.0 had adopted guidelines of Section 508 § 1194.22, a standard for accessible web-based intranet and internet information and applications. Similar to Section 508 § 1194.22, KWCAG1.0 puts into level 1 guidelines of WCAG1.0, a W3C prepared guidelines' priority 1 accessibility checkpoints. Satisfying the priority 1 checkpoints of Web content accessibility guidelines 1.0 is a basic requirement for some groups to be able to use web documents, that a web content developer must satisfy these guidelines.

3 Accessibility Evaluation

3.1 Guideline 1. Provide a Text Equivalent for Every Non-text Element

We checked how well the e-Governmet web sites comply to the guideline 1 of the WCAG 1.0. The use of "alt" and "longdess" was checked for all the graphic elements of web pages. These include all the images, graphical representations of texts, image map regions, and applets.

Overall, only 10% of the non-text element of the sites were provided with a text equivalent. One of the site had as much as 63% of compliance, while many sites did provide none for the text equivalent to the non-text elements.

This is a very significant problem, in that Korean web sites use lots of graphical images, pictures and graphical representations of texts.

3.2 Guideline 2. Don't Rely on Color Alone

We manually checked the color coded information on the e-Government web sites to examine the compliance to the guideline 2. None of the color-coded information provided another way to convey the information along with the color. Even though the manual checking was a very time consuming process and entailed subjective judgment, non-compliance was so apparent.

3.3 Guideline 9. Design for Device-Independence

We manually checked for device-independency of the selected sites. We checked if the sites can be navigated by only the keyboard without using mouse on these sites. The result was only four out of the seven sites were device independent. This means that the other three sites cannot be accessed without a mouse or equivalent pointing devices. We did not include any of the voice activation.

3.4 Guideline 12. Provide Context and Orientation Information

We checked how well the e-Government web sites comply to the guideline 12 of the WCAG 1.0. The use of title attributes on frames was checked. To use the "title" attributes on each frame is needed to facilitate frame identification and navigation.

Overall, only 9% of the frames of the sites were provided with a title attribute on the frames. One of the site had as much as 30% of compliance, while many sites did provide none for the title on frames. However, about 80% of the pages were provided with a title attribute, which makes it easier for screen readers to provide the page description.

4 Conclusion

This study evaluated the web sites that consist of e-government of Korea. Even though we evaluated selected sites of the e-Government of Korea, the result showed that there was a lack in accessibility even in some major checkpoints of KWCAG 1.0, such as providing text equivalents to the non-text elements. As was reported in the

Brown University's global e-Government evaluation report, Korean e-Government provided only approximately 10% of the checkpoint 1 accessibility features for the selected sites.

Even though it was reported that the web sites of Korean government's ministries and agencies complies about 34% of the checkpoint 1 accessibility features [1], the accessibility of the e-Government was at a lower level. This means that lack of coherent government policy for web accessibility causes lack of providing standard e-Government services for the part of the population with special needs. Government is required to set a policy for web accessibility together with an enforcement procedure for all the government web sites including e-Government. Results also suggested that the accessibility among the different sites was not evenly distributed. This is probably due to the fact the e-Government portal was simply acting as a mere gate to all the ministries' civil affairs service sites, not being a well-structured portal that should be organized and pre-planned as a whole. It should also be noted that without specific government policy or guidelines that dictate the accessibility concern, the design tends to focus more on aesthetics and efficiency by young designers of the private design companies who are not very well aware of the notion of accessibility. The inconsistent distribution of accessibility among the services that we observed during the tests also can be attributed to the lack of consistent guidelines that run through all the government ministries and agencies.

In general, governments need to understand the obstacles to making e-government's web sites accessible and should adopt the appropriate solution to improve it. They have to spread the awareness of the importance of accessible e-Government sites by developing appropriate policies. It is also important to understand that all e-Government endeavors are critically dependent on the accessibility of its integral portal [1]. The civil affairs services provided by e-Government are different from the individual government agency web sites in that actual transactions for civil services are available while the government web sites are only for providing information on the administrative agenda and public relations of the agencies. Therefore, e-government transactions have to be available for 24 hours a day, 7 days a week, particularly for those citizens who have handicaps.

Acknowledgments. This research was supported by the MIC(Ministry of Information and Communication), Korea, under the ITRC(Information Technology Research Center) support program supervised by the IITA(Institute of Information Technology Assessment), IITA-2006-(C1090-0603-0046).

References

1. Abanumy, A., Al-Badi, A., Mayhew, P.: e-Government Website Accessibility: In-DepthEvaluation of Saudi Arabia and Oman. Electronic J. e-Government 3(3), 99–106 (2005)
2. Joonho, H., Doojin, C., Sukil, K: An active step toward a web content accessible society, International Cross-Disciplinary Workshop on Web Accessibility(W4A Workshop) at WWW2005 Conference, Chiba, Japan (2005)

3. Information Society Commission: eGovernment: More Than an Automation of Government Services (October 2003)
4. Seifert, J., Relyea, H.: Considering E-Government from the U.S. Federal Perspective: An Evolving Concept, A developing Practice. J. E-Government 1(1), 7–14 (2004)
5. Vassilakis, C., Lepouras, G., Fraser, J., Haston, S., Georgiadis, P.: Barriers to ElectronicService Development. E-Service Journal, pp. 41–63 (2005)
6. West, D.M.: Global E-Government 2006 Brown University (2006)

The Potential of Web Accessibility in China: A Hypothesis on Its Impact on the Global Web Interface

Eleanor Lisney, Christina Li, and Songtao Liu

215-1-101 Cuicheng, Fatou, Chaoyang District, Beijing, China
elisney@gmail.com, christina.li@uigarden.net,
sean.liu@uigarden.net

Abstract. This paper is based on a preliminary research conducted in 2005 into the current level of web accessibility in China. We evaluated a few websites which are meant to be information portals for the visually impaired. Results of the evaluations show none of them is compliant with Web Content Accessibility Guidelines 1.0 (WCAG 1) [1] Level A checkpoints and most of them are inaccessible at all. Causes behind this situation may include lack of legislation from the government and lack of accessibility awareness of web designers. However, we found Chinese Web accessibility also faces more of a challenge as a result of the complications of Chinese language itself. We hope this paper could bring the attention of both Chinese government and Chinese web designers on web accessibility and also help international stakeholders to understand the facts in China.

Keywords: web accessibility, China, visually impaired, language, universal design, usability, Chinese.

1 Introduction

This paper is based on a preliminary research into the current level of web accessibility in China. Web accessibility has been given a priority status in the west as a result of legislation such as Section 508 of the Rehabilitation Act in the United States and the Disability Discrimination Act (DDA) in the United Kingdom. In China, accessibility for people with disabilities in itself is not a well-known concept. Web accessibility is practically an unknown concept except perhaps, for those who use assistive technology such as text readers or browsers to access the web. And even then, it is an extremely limited access; very, very few Chinese websites take web accessibility into consideration in the design.

In this report, we show that even websites which are meant to be information portals for the visually impaired are not compliant with Web Content Accessibility Guidelines 1.0 (WCAG 1) Level A checkpoints. We also look at the barriers of different dialects and languages used in China

This article is a collaboration of Chinese usability specialists from across the globe to report on the current status of web accessibility in China: are there any developments or efforts towards inclusive design? Events such as the recent User Friendly 2006 Conference held in Beijing show that there is awareness that usability

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 79–87, 2007.
© Springer-Verlag Berlin Heidelberg 2007

is an important process of web design. Ongoing research is being conducted and this article is a result of the initial findings.

2 A Survey on Chinese Web Accessibility

In 2005, we conducted an evaluation of web accessibility situation on the homepages of some websites in Chinese. The objectives of this evaluation are 1) to check if these web pages are WCAG 1.0 compliant; 2) to identify the common accessibility issues in these web pages if there are any; 3) to analyze the findings.

These websites are mainly for disabled people, especially the visually impaired people in China. They are homepages of:

China Disabled Persons' Federation (CDPF) website;
Blind Online Channel on CDPF website;
CDPF website, English Edition;
China Braille Press (CBP) website;
Yongde Website for the Blind;
China Massotherapy Network;
Huaxia Website for the Blind;
The Hadley School for the Blind in Fuzhou China.

Table 1. A brief description of the organizations that host the websites

Name of Organization and URL	Organization Description
China Disabled Persons' Federation (CDPF) http://www.cdpf.org.cn	CDPF is the unified organization of and for over 60 million persons with various categories of disabilities in China
CDPF - Blind Online Channel http://www.cdpf.org.cn/mrzx/index.htm	A CDPF website Channel for the blind people
CDPF, English Edition http://www.cdpf.org.cn/english/index.htm	The English Edition of CDPF website
China Braille Press http://www.cn-bp.com	China Braille Press (CBP) is a branch organization under CDPF. It is also the only Braille press house in China
Yongde Screen-reader www.wangyongde.com	The website for Yongde Screen-reader, one of the most popular Chinese screen-reader for the visually impaired. Wang Yongde, himself a blind, produced the software.
China Massotherapy Network for the Blind http://www.bjanmo.com/	Massage Therapy is the most common and most possible job that the blind people in China can have. This website is mainly for the blind massage therapist.
Huaxia Website for the Blind http://www.chinamr.org/	A non-profit website for the blind community in China.
The Hadley School for the Blind in Fuzhou China http://www.hadley-china.org/index2.htm	The Hadley School for the Blind in Fuzhou, China

These websites are selected because they are mainly for the visually impaired people and they are meant to be compliant with accessibility standard and guidelines. Table 1 depicts what these websites are about.

2.1 Methodology

We use the automated Bobby Online Portal [2] combined with manual heuristic evaluation to examine if these web pages are compliant with WCAG 1.0.

First, we use Bobby Online Portal. For each web page, we input the URL into Bobby Online Portal to generate a report. Bobby Online Portal provides reports based on Section 508 and WCAG 1.0 Guidelines. In our evaluation process, we check WCAG 1.0, Priority 1 only. Second, we heuristically and manually check some triggered issues that Bobby cannot determine automatically.

If a web page passed such a Bobby test, it means this web page qualifies as Bobby A Approved, which means it is accessible in the Support Level A of WAI Content Accessibility Guidelines.

We realize that automated testing can detect only some of the accessibility violations identified by WCAG 1.0 Guidelines. It provides an initial assessment of the technical accessibility of a website. However, an automated tool cannot check on color contrast or evaluate the adequacy of the alt text description. For the purpose of this paper, we only consider the homepage. We used Watchfire's Bobby tool rather than others because it is a widely known tool, more so than Cynthia or WAVE, for example. We did not expect to find and nor did we find features such as 'Skip navigation links' nor were there any style sheets associated to separate presentation from content.

2.2 Results of Bobby Evaluation

First we use Bobby Online Portal. The summary of the Bobby test results are listed in Table 2 below.

Table 2. Results of web accessibility evaluations on the selected websites

Name of Organization and URL	Bobby A Approved?	Major Accessibility Error(s)
China Disabled Persons' Federation (CDPF) http://www.cdpf.org.cn	No	Provide alternative text for all images. (117 errors)
CDPF - Blind Online http://www.cdpf.org.cn/mrzx/index.htm	No	Provide alternative text for all images. (3 errors)
CDPF, English Edition http://www.cdpf.org.cn/english/index.htm	No	Provide alternative text for all images. (17 errors)

Table 2. (*Continued*)

China Braille Press http://www.cn-bp.com	No	Provide alternative text for all images. (1 errors)
Yongde Screen-reader www.wangyongde.com	No	Provide alternative text for all images. (15 errors)
China Massotherapy Network for the Blind http://www.bjanmo.com/	No	Provide alternative text for all images. (5 errors)
Huaxia Website for the Blind http://www.chinamr.org/		Provide alternative text for all images. (3 errors)
The Hadley School for the Blind in Fuzhou China http://www.hadley-china.org/index2.htm	No.	Provide alternative text for all images. (3 errors)

No homepage of these 8 websites is qualified as Bobby A Approved. All web pages investigated have one issue in common: they don't provide alternative text for all images.

As the automated Bobby online Portal is not accurate enough to evaluate a web page's accessibility, to compensate this, we do manual heuristic evaluation in the next step.

2.3 Results of Manual Heuristic Evaluation

During our manual heuristic evaluation, we notice that some of the 8 investigated homepages, though they are not qualified as A Approved by Bobby, but they are still reasonably web accessible. These websites are:

China Braille Press
The Hadley School for the Blind in Fuzhou China

We find the images with no alternative text descriptions are not relevant - they contain no information at all. They did not use empty alt tags as is usually recommended by WCAG 1.

The home page of CDPF comes out the worst in this comparison. There are 117 instances of images without alternative text descriptions. If we take a look at the screenshot of that home page, we can also identify some 5 flicker elements that could make accessibility even worse. There is a pop-up window, which is a sheer image with no alternative text description.

The result of evaluation on Yongde Screen-reader software homepage is negative too. Yongde screen-reader is one of the most popular screen-reader in China. Its website should be accessible. But again, many button images that convey important information for visitor to click on have no alternative text descriptions (Figure 3).

Fig. 1. CDPF homepage screenshot. There are flickers (framed in black shown in the figure), a pop-up window, and 117 images without alternative text descriptions, which is not accessible for those who rely on screen-readers.

Fig. 2. Screenshot of CDPF English edition homepage. Images that convey important information, as framed in red, have no alternative text descriptions.

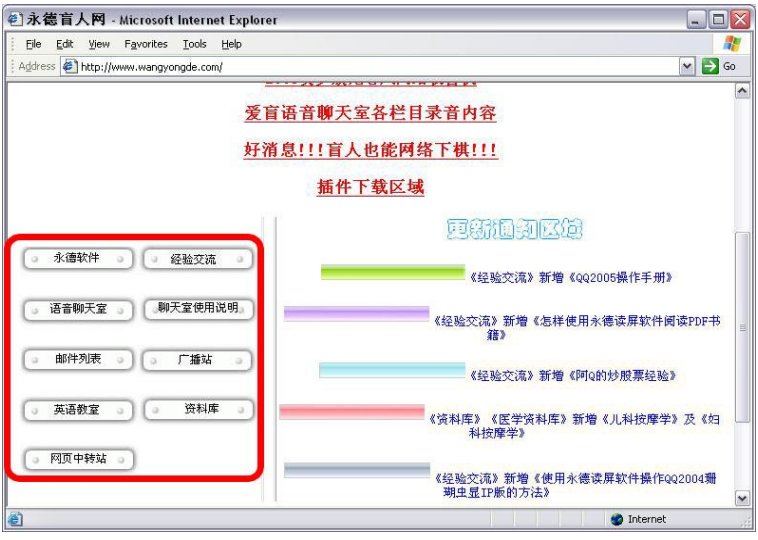


Fig. 3. The important button images on the left (framed in red) that provide navigation function have no alternative text descriptions

To summarize the findings, we rate these 8 websites with 'Accessible', 'Partially Accessible', and 'Not Accessible' as follows (Table 3):

Table 3. Evaluation results – Our rating

Name of Organization and URL	Bobby A Approved?	Our Ratings
China Disabled Persons' Federation (CDPF) http://www.cdpf.org.cn	No	Not Accessible
CDPF - Blind Online http://www.cdpf.org.cn/mrzx/index.htm	No	Partially Accessible
CDPF, English Edition http://www.cdpf.org.cn/english/index.htm	No	Not Accessible
China Braille Press http://www.cn-bp.com	No	Accessible
Yongde Screen-reader www.wangyongde.com	No	Not Accessible
China Massotherapy Network for the Blind http://www.bjanmo.com/	No	Partially Accessible
Huaxia Website for the Blind http://www.chinamr.org/		Partially Accessible
The Hadley School for the Blind in Fuzhou China http://www.hadley-china.org/index2.htm	No.	Accessible

2.4 Results

As we can see from Table 3, 2 out of 8 homepages are accessible based on our investigation:

China Braille Press
The Hadley School in Fuzhou, China

3 of them are rated as 'Not Accessible':

CDPF website
CDPF English Edition
Yongde Screen Reader website

The results are revealing in themselves, as websites that serve the blind community are expected to be accessible. But that is the real situation in China. There should be more advocacies for disability awareness and web accessibility.

3 Language, Dialects and Other Barriers

Even though a website itself is qualified as accessible, for the visually impaired, technology such as a screen reader is still needed to aid to access it. The screen reader software translates texts to speech so the visually impaired can access web pages.

However, the Chinese language presents special challenges for software including text readers. Chinese language is made of a combination of characters and sounds. Each character is more like a drawing rather than a letter. The Chinese text consists of a string of characters and each character is to be pronounced separately. A Chinese word may contain one or more characters assembled to make another word. A text reader needs to be able to recognize the various permutations of a Chinese text string in the lexical context to produce the correct pronunciation. Because the Chinese language uses a logographic script—one in which one "character" corresponds roughly to one "word" or meaning—there are vastly more characters, or glyphs, than there are keys on a standard computer keyboard. A variety of keyboard input method editors or IMEs have been designed to allow the input of Chinese characters using standard keyboards. This added complication of encoding of the Chinese in websites can prove to be too much of a challenge to the development teams.

There are sophisticated methods for guessing which characters the user intends based on the context of the sentence. For the blind user, there is the added problem that the transfer of Chinese texts into Braille is not 100% accurate. Older Chinese texts might be only about 80% accurate; because the input into Braille uses pinyin. The text reader has also to accommodate English as many of the links, URLS and software are in English.

The language complications that adds to the difficulties of consolidating and working together to form a single text reader that would work across board in China is compounded by the fact that China has as many dialects again. Putonghua or Mandarin is the national language and it is taught in schools but local dialects still prevail. For example Cantonese is still the de facto dialect in Guangzhou and in Hong

Kong and they have their own text readers. A Beijing resident would not be able to understand a Cantonese in his mother tongue but they would have no problem with the written language if they can read it.

4 Challenges and Potentials

As we have shown, the Chinese blind user faces more of a challenge as a result of the complications of Chinese language itself. This means whatever few advances and benefits have made accessible the Western world on the World Wide Web cannot be assimilated into the Chinese world. WCAG 2 is apparently going to be less Anglo centric but it still has not much on non-European languages.

On the other hand, those who have been trained and have access to computer equipment are certainly appreciative of being able to communicate electronically and have access to information on the Internet. Wenru Niu shows from her report that the schools seem to on their way to teach their students computer literacy that might eventually offer them better prospects for work other than the traditional massage jobs. At the moment they are struggling with the needs for better curriculum, teacher training to teach the students and up to date computer equipment. The importance of training for the teachers is not to be underrated. As anybody who has used any of the text readers could easily emphasize, learning here is not readily available in a do-it-yourself kit from a manual, especially when you have not learned to read. Students from the Qingdao School for the Blind are adamant that "they need a screen reader that will work in programs that sighted use. They do not need a special program that only works for the blind." [3]

Given the momentum of the growing web usage in China, web developers have become aware of usability issues and the need to conform to world standards. Some of the new technology can be of benefit to both the blind and the sighted world such as the Tongchuang Blue Sky software from Harbin (of which there is not much information) and CU VOCAL from Hong Kong. Universal design is still the best way forward. Assistive technology always carries a heavy price for those who can least afford it and there is the possibility to work with open source and free ware. However, it is not as prevalent and available as commercial software.

Some of the legislation passed in the previous year shows that China is taking account of the discrepancy and is doing something about it. In the year of 2004 and 2005, we saw twice'China Information Accessibility Forum' conferences held in Beijing China, hosted by CDPF and China Association of the Blind (CAB). This is the first and biggest ever event in China to advocate and promote accessibility issues for the blind. The China Committee of Information-without-Obstacles for the Deaf followed it on November 2005. One aim of those conferences is to advocate the information accessibility from technology aspect and the government.

5 Conclusion

We conducted web accessibility evaluation on some selected websites which are dedicated for the people who need accessibility most. We found only 2 out of 8 are

accessible, and none of them passed the WCAG 1.0 Level A checkpoints. Obstacles include lack of legislation and lack of accessibility awareness from web designers. Difficulties from the Chinese language itself exist and should be considered in the future standard or guidelines if Chinese government will have. Japanese guideline can be referred as the Japanese language uses Chinese characters everywhere.

We propose a more extensive survey on web accessibility in China be carried out in future that can cover more websites in a broader range. User testing of disabled people including blind people and partially sighted people can be helpful to understand more.

Why should Chinese web accessibility matter to technical communicators in the international community? Apart from the ethical issues for pursuing inclusive design, Chinese web developers are becoming strong players on the World Wide Web, it is important that the newly developing market should have the essential dimension of technical web accessibility. A commitment to an accessible information architecture that is usable from the initial stages of development seems logical before huge investitures get entrenched in large websites which are not accessible. As the Chinese seem poised to become a major power in the technical market place [4], technical communicators would do well to keep abreast of events there for opportunities. At the moment, technical communicators might want to look at future trends such as the wireless technology that is growing by leaps and bounds there. In the meanwhile, technical communicators who plan to work or have communication with the Chinese should understand the level of web accessibility there and encourage inclusive design because ultimately the international community would benefit in such an exchange. One of the initiatives that facilitates just such an exchange is the uiGarden.net [5], a bilingual webzine which offers an East West interaction and to bridge between the different cultures in the usability field.

References

1. Web Content Accessibility Guidelines 1.0 http://www.w3.org/TR/WAI-WEBCONTENT/
2. Bobby Online Portal http://bobby.watchfire.com/bobby/html/en/index.jsp
3. Wenru, N.: CHINA TRIP REPORT October 9-November 9, 2003 part I: unpublished report (2003)
4. St. Aman, K: Considering China. A Perspective for Technical Communicators Technical Communication 48(4), 385–388 (2001)
5. UIGarden Website, EastWest Interaction http://www.uigarden.net

Restoring Semantics to BML Content for Data Broadcasting Accessibility

Kinji Matsumura, Yasuaki Kanatsugu, Takuya Handa, and Tadahiro Sakai

NHK Science & Technical Research Laboratories,
1-10-11 Kinuta, Setagaya-ku, Tokyo, Japan
{matsumura.k-ke,kanatsugu.y-ki,handa.t-es,
sakai.t-jk}@nhk.or.jp

Abstract. This paper proposes a method to enable visually impaired viewers to access data broadcasting content by using various accessible devices. Data broadcasting, coded in Broadcast Markup Language, has been almost inaccessible because of its usage of dynamic HTML features and lack of semantics. This method uses a template rule to transcode BML content into an accessible static description with structure and semantics. An implementation of a trial receiver system and an experiment on transcoding are described.

Keywords: Data Broadcasting, BML, Accessibility, Transcoding.

1 Introduction

Digital data broadcasting started in Japan in 2000. Although its content is coded in Broadcast Markup Language (BML)[1], no accessibility supports have been implemented so far. To improve the accessibility of digital broadcasting, we developed a method to transcode BML content into a static and structural form with sufficient semantics to make it accessible to persons with visual impairments.

The goal of this research is to develop a unified accessible platform for presenting various content types on different kinds of devices for visually impaired users, by defining and utilizing a common set of APIs and metadata. In line with this concept, this paper explains our development of a transcoding method to make BML accessible on the unified platform.

2 Background

2.1 Research Framework

The goal of our development is to realize a unified platform that provides visually impaired people with accessibility to various multimedia contents such as HTML, Flash, and data broadcasting, by using user interface methods such as voice browsers, tactile displays, and Braille displays. As the key feature, we aim to define a common set of APIs and metadata. Fig. 1 shows an overview of the unified platform and an

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 88–97, 2007.
© Springer-Verlag Berlin Heidelberg 2007

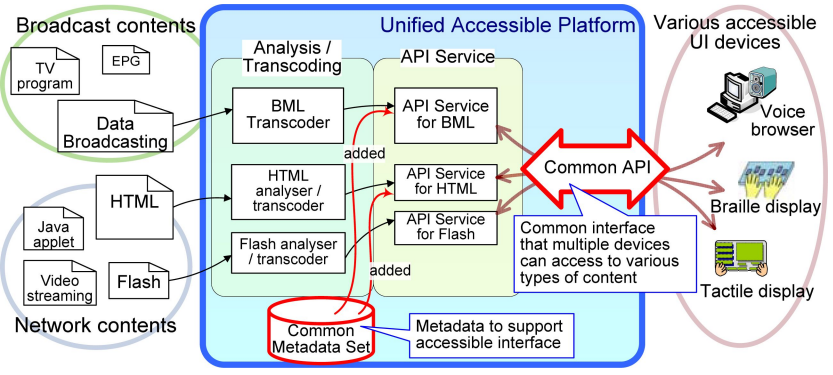


Fig. 1. Overview of unified platform

example of how different kinds of contents are provided through the common API. By using a device or software that complies with the common API, users can access different kinds of multimedia content through a uniform procedure.

2.2 Present Situation of Accessibility in Data Broadcasting Services

Data broadcasting provides a variety of information by using BML. Although several accessibility supports for TV programs such as closed captions, sub-channel commentary and speech rate conversion have been developed so far, no such support has been implemented in data broadcasting, which means that visually impaired people do not have access to data broadcasting content.

Table 1. Composition of BML specifications

XHTML 1.0	Basic structure and declaration of text and visual objects[1]
CSS level 1/2	Graphical presentation design[2]
ECMAScript	Data processing and dynamic interaction
DOM Level 1/2	Dynamic interaction
Enhancements in ARIB STD B24	Broadcast specific functionalities: Navigation with remote controller, AV stream/event handling, and receiver functionalities such as Rec. reservation, persistent memory read/write, etc.

As shown in Table 1, BML is based on Dynamic HTML technology, and it has enhanced functionalities for broadcast use, which is prescribed in ARIB STD-B24 [1]. In an ongoing service, though, the user interface design is restricted under the premise that content is presented only on a TV screen and users interact with the content by using a remote equipped with four-way arrow keys and an OK button. For this reason, the available XHTML tags and CSS properties are limited in actual operations.

[1] Available tags are strictly limited in actual operations.
[2] A fixed layout using absolute positioning is required in actual operations.

3 Requirements Analysis

3.1 Requirements of User Interface and Basic Concept of Common API

As the first step of our study, we derived the fundamental functionalities that the platform should provide via the common API. We assessed the requirements of visually impaired users in regard to accessing general multimedia content with various user interface methods, particularly considering devices with different characteristics such as Braille displays [2] and 2D tactile displays [3] as well as ordinary voice browsers. The requirements are as follows:

- *Content should be fragmented into information units, and the units should form a tree or a table structure reflecting the content's essential structure.* Content should be provided in a versatile form in order to allow various devices to present content in an appropriate style utilizing the device's characteristics and abilities. Hence, the common API should be designed to provide information in a form whereby content is fragmented into information units and is structured according to the content's original semantic structure or navigational structure. Also, the structure should be static with no dynamic alterations, to ensure that the user understands it.
- *Semantics should be attached to each content fragment.* Semantics such as meaning and role should be added to let devices use them in order to behave correctly. Existing metadata such as [4] should be utilized.
- *Alternative or descriptive annotations should be added.* An alternative textual description should be added to visual content or to long and complicated content.

Based on these fundamental requirements, we defined the basic concept of a common API. The API is a simple set of commands to provide content according to the content's tree/table structure with semantic and alternative descriptions. The user's devices call these APIs to navigate through the content, acquire information fragment, and then present the content in a manner appropriate for the user agent's characteristics. Note that we do not describe the detailed specifications of the API in this paper.

On the other hand, this concept becomes a functional requirement for the transcoding process (see Fig. 1 – analysis / transcoding layer) that makes multimedia content accessible via the common API. Such a process depends on each content format. Many studies have been done on such technologies for WWW content formats.

3.2 Technical Issues of BML Content

We analyzed the characteristics of actual broadcast BML contents to determine issues affecting accessibility. We found the following problems specific to BML:

A. *Very poor structural and semantic information due to tag usage restriction.* As mentioned above, BML content has very little structural and semantic information. Tags are mostly used for declaration and layout of visual objects. Therefore it is difficult to extract structural and semantic information from tags.

B. *Most content is transmitted in an external data file.* Most of the essential information is stored and transmitted in an external data file named Binary Table. A BML document file only has a description of how it is presented, such as CSS layouting and ECMAScript code for dynamic UI behavior and data processing. Also, in many cases, the data in the Binary Table are processed in ECMAScript before being presented. As shown in Fig. 2, for example, in order to present personalized weather forecast, the script code searches a Binary Table by referring to the user's preset postal code, and then the retrieved value is used as a reference to display an image that represents the corresponding forecast.

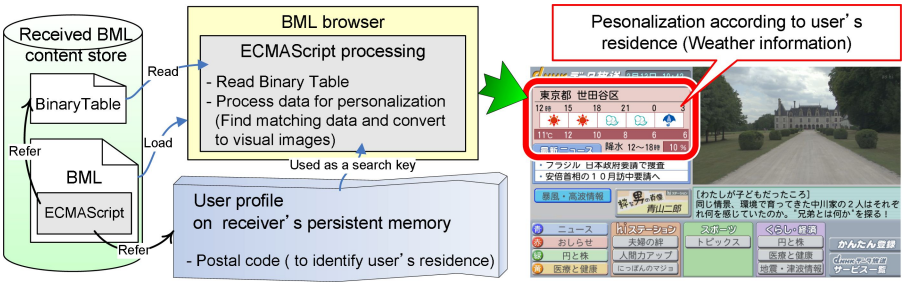


Fig. 2. Example of Binary Table data processing

C. *Input event handling specifically designed for remote controller UI.* In DHTML and Flash, user-input events are usually linked to corresponding visual objects on the screen, so the target of the user's operation can be identified by observing these objects. But in several BML contents, user inputs from a remote are directly processed with a script code. In such cases, an object that looks visually focused does not correspond to the browser's internal focused state. For example, in a content shown in Fig. 2, all 'onkeydown' events are obtained by only one div element and script processing handles visual focus-moving presentation.

D. *Heavy use of custom GUI – especially for page-flip GUI.* Custom GUIs such as tab and pop-up menu are used. Notably, page-flip GUIs (Fig. 3) are used in many cases. In a page-flip GUI, temporal fragmentation of content on the same DOM tree node prevents content analysis and structuring.

These characteristics also lead to an increase in the size of script code in BML. According to our research, in more than 130 of the BML documents we looked at, 66.6% of the lines are ECMAScript code. This also shows that most BML documents have characteristics that make it harder to improve accessibility.

Our analysis indicated that BML has characteristics similar to dynamic content format such as DHTML, Flash and Java applets, despite it being based on XHTML. Moreover, it is sometimes much more complex because its design is specific to a remote controller UI.

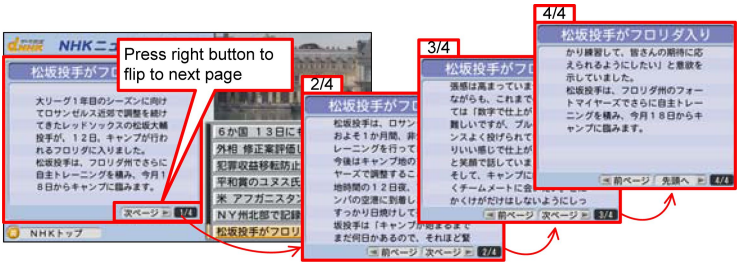


Fig. 3. Example screenshot of page-flip GUI

3.3 Broadcast Service Requirements

This section outlines the requirements specific to the broadcast service.

- *Information must be correct.* A broadcast service must provide accurate information. The possibility of transcoding errors should thus be minimized.
- *The impact on existing receivers should be avoided.* Broadcast systems are huge systems with millions of anonymous receivers. Therefore, when a new additional service is implemented, the compatibility with existing receivers must be maintained.

4 Study on Possible Solutions

We used the results of the aforementioned study to determine a feasible method to transcode BML content. We compared five possible methodologies by using the following evaluation criteria: a) quality of an expected service, b) cost for additional authoring, and c) impact on existing receivers. Criteria a) and c) are derived from broadcast service requirements and we also considered criterion b) because it is an important factor in operating actual services. We evaluated the five methodologies as follows:

1. *Static analysis and annotation for BML document.* Many of existing accessibility technologies adopt this methodology. Tags or the visual layout of content are analyzed statically and additional metadata annotations such as described in [4] are used to add semantics.
   Evaluation: a) Additional annotations can help to alleviate the problem of BML's lack of semantics. However, most BML content will not be accessible since the static analysis does not support dynamic content. b) Cost is relatively small, although it depends on the amount of additional annotations that must be prepared for each content. c) The impact is avoidable if annotations are separated from the BML content to avoid causing malfunctions in existing receivers. We also must consider how to distribute annotations.
2. *Observing the state of the browser's internal state.* In addition to the techniques described in i, the content player software (i.e. BML browser) sends its internal DOM tree state so that analyzer/transcoder software can obtain information on

dynamic content changes. The accessibility API of MSAA[5] adopts this methodology.
Evaluation: a) Observing changes in the DOM tree can solve some of the problems caused by dynamic content. However, BML has characteristics that prevent this methodology from working. In cases like 3.2 C, the user's target can not be obtained by checking the DOM tree, so the analyzer/transcoder software cannot identify which object is being manipulated. Also, in cases like 3.2 D, the acquirable data is fragmented and temporally separated. Hence, it is hard to determine the order of the fragments, and thus, content is difficult to structure. b) Almost the same as i. c) The same as i.

3. *Embedding script code for accessibility purposes.* In this methodology, a content author embeds script code for accessibility purposes. The embedded code describes the procedures so that the accessible devices can behave in a similar fashion as designed for the normal presentation devices. This methodology is generally used in combination with the techniques described in ii (see[6]).
   Evaluation: a) If all the necessary descriptions are properly coded, users can interact in almost the same manner as the content is designed for general viewers. b) The amount of code tends to increase accordingly. Because the BML document, by its nature, has complex script code, the cost for authoring and maintenance may be a serious problem. c) It is difficult to separate the original code from the additional code for accessibility. To avoid conflicts with existing receivers, conditional statements to determine the receiver's supported functionality must be written everywhere. This also leads to a cost problem.
4. *Static analysis and annotation of Binary Table.* As a new approach, we considered a methodology that applies external static annotations to a Binary Table to make it structured. Data in the Binary Table are directly used for the presentation without using a script code in the BML.
   Evaluation: a) Static annotation does not work when Binary Table data are processed before presentation. b) Binary Table has no structural information, so the annotations must describe the structure from scratch. c) Similar to i.
5. *Applying an external template with data processing to the Binary Table.* We also considered the methodology that extracts data from the Binary Table and processes it before running it down to the structured template. This does not use the original BML document either, but it is different from methodology iv in that data processing is possible.
   Evaluation: a) By using a properly authored template, it can produce similar content to the original BML because data processings such as search and text

Table 2. Comparison of possible solutions

	a) Quality	*b) Cost*	*c) Impact*
i. Static annotation to BML document	Bad	Good	Fair
ii. Dynamic observation	Fair	Good	Fair
iii. Embedded accessibility script code	Good	Bad	Bad
iv. Static annotation to BinaryTable	Bad	Fair	Fair
v. Process and structure BinaryTable	Good	Poor	Fair

processing are possible. b) Relatively high because template author must code both the structure and data processing from scratch. However, it does not have a complexity problem like that of iii. c) Almost the same as methodology i.

Table 2 summarizes the above analysis. To give a quality of service close to that of existing broadcast services, we decided that criterion a) is the most important. As a result, we chose methodology **v** for our system because only **v** meets the required quality level and has sufficient feasibility. We believe that the cost problem can be handled because the amount of broadcast content is much smaller than on the Internet.

5 Design of Transcoding Method

5.1 Template-Engine Mechanism

We designed a concrete BML transcoding scheme by adopting methodology **v**. Fig. 4 shows the sequence of transcoding. In this model, documents are transcoded according to template rules that have descriptions for data extraction, processing, and structuring. The transcoder finds and applies a matching template rule to the original BML document, and generates similar content to the original one but it is static and structured. We provisionally named the coding scheme of trancoded content Simplified BML (SBML). The SBML processor processes transcoded documents, then provides information to devices via the common API.

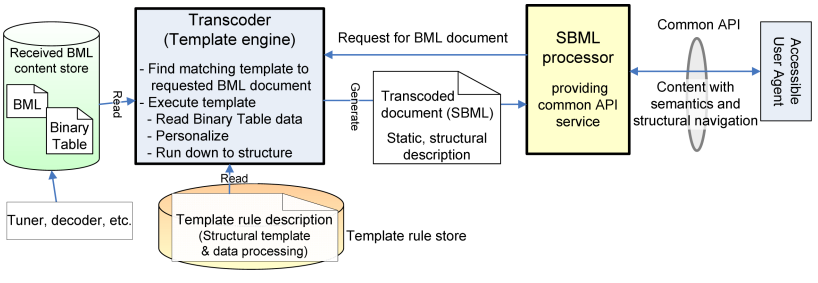


Fig. 4. Accessible BML browsing system with transcoding mechanism

The transcoding process of this model is similar to existing technologies of template engines such as PHP and JSP, although it is different in that it runs on the receiver side. Thus, in the prototyping, we used Ruby scripting language (erb) for the template description and processing.

5.2 Consideration of Authoring and Distribution of Template Rules

As with existing annotation-based technologies, these template rules depend on the original content and they must be prepared for each original BML content. To be more precise, template rules depend on the information structure of the original content, i.e., the Binary Table's data scheme and the usage of its data.

On the other hand, most BML content is automatically updated using authoring templates. Therefore, template rules can be defined by making use of the design pattern of the original authoring templates. Fig. 5 shows the proposed distribution system. By distributing these rules to receivers beforehand, receivers can transcode updating content at any time by receiving only broadcast waves. A matching template rule can be identified by the received BML's broadcast URL. When the authoring template is replaced, however, the corresponding new template rule must be redistributed by using the network or other means. The other advantage is that the existing data broadcasting system is not affected at all.

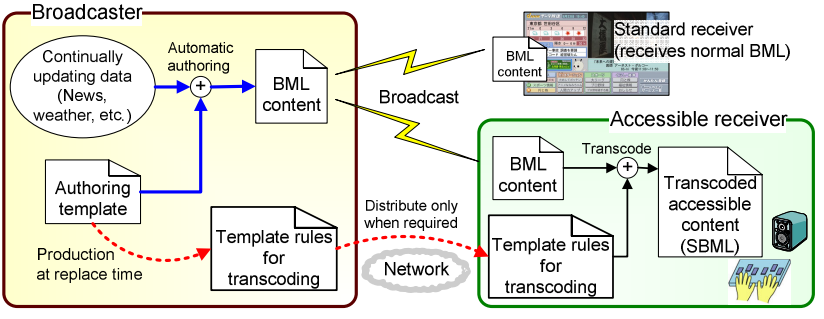


Fig. 5. Broadcast system with template rule distribution for content transcoding

5.3 Description Scheme of Transcoded Content

We also defined the description scheme of the transcoded content – SBML. The scheme must be able to describe structure and semantics. On the other hand, enhanced functionalities defined in the BML specification (such as access to persistent memory and broadcast event handling) should be available as long as the script code does not dynamically alter the DOM tree structure of the content.

For these reasons, we determined that the description scheme should be based on XHTML (including structuring tag sets) with BML enhancements, but without

Table 3. Differences between BML and SBML

	BML (currently operated)	*SBML (provisional) (for accessibility use)*
Document Structure	Mostly only p, div, object	Same as on the left. Plus, h1-h6, tables, forms, etc.
DHTML feature	Used	Not used
Presentation design	Strictly specified in content	Not specified in content
Broadcast specific APIs & tags	Available	Available
Metadata for accessible user agent	Not available	Included as compound document

DHTML features. In addition, metadata for accessibility is available to form a compound document. This can be regarded as another profile of BML, in contrast with the one used in actual operations. Table 3 shows the differences between SBML and current BML.

6 System Prototyping

We designed and implemented a prototype receiver system with the functionalities listed below:

- Receives BML contents from broadcast waves and transcodes them in real time.
- Applies corresponding template rules to each BML content.
- Provides common API service processing transcoded SBML. (The provisional specification was used for the common API to meet the requirements described in 3.1.)

We used the system to confirm the feasibility of the transcoding method. The experiment used NHK's actual broadcast content and we authored template rules for 31 BML documents, all of which generated SBML documents as we expected. Below is example code for a template rule. Codes between <% and %> are descriptions of data processing, and the rest are templates that describe structure and semantics.

```
<?xml version="1.0" encoding="euc-jp"?>
<?bml bml-version="_._"?>
<html>
<head><title>NHK Anytime News</title></head>
<body style="clut:url(news.clt)">
<%  f = BinaryTable.open("~/0011/news.btb","1,S:1V")
    rowData = f.getRowData()
    /(....-..-..-)(..)\.(..)\...\...../ =~ rowData[0]  %>
<object type="image/png" data="NHK.png" alt="anytime news" />
<%= $2 %>:<%= $3 %> updated
<%  while rowData = f.getRowData()
      if (rowData[0]=="newsreportstartmarker") %>
      <h1 id="news<%= I %>"><%= f.getRowData() %></h1>
<%  else %><%= rowData %><%
   end
end %>
<a accesskey="Y" href="/40/0000/top.bml">NHK Top Page</a>
</body>
<html>
```

We confirmed that this system can transcode received BML content into an accessible form and provide the content to devices via the common API. Also, by receiving BML contents at different times, we confirmed that a single template rule can be applied to different versions of content that are generated from the same authoring template.

7 Conclusion

We described the development of a method to make BML content accessible to visually impaired people. We analyzed the requirements and found that BML has specific problems affecting accessibility. We proposed a static pre-transcoding methodology for BML and implemented a trial system to confirm its feasibility.

We will evaluate the quality and precision of the proposed transcoding method in the near future, as well as provide a justification for the common API design. The BML transcoding method proposed in this paper can be utilized for other purposes other than improving accessibility. By transcoding BML into static and device-independent desciptions, the possibilities for versatile usage of data broadcast content will be greatly enlarged.

As a system functionality, we are planning to enhance the system so it will be able to handle broadcast events triggered from broadcasters at an asynchronous timing.

Acknowledgements. Part of this research was carried out with the support of the National Institute of Information and Communications Technology.

References

1. ARIB. ARIB STD-B24, Data Coding And Transmission Specification For Digital Broadcasting (Version 5.0) (2006) http://www.dibeg.org/aribstd/ARIBSTD.htm
2. Sakai, T., et al.: Development of an Integrated Information Receiving Terminal for Visually Impaired Persons. ICCC 2,I-2-06 172–179 (1999)
3. Sakai, T. et al.: Improving Access to Digital Broadcasting for Visually Impaired Users, HCI International (2005)
4. W3C: Role Taxonomy for Accessible Adaptable Applications Working Draft http://www.w3.org/WAI/PF/GUI/roleTaxonomy-20060508.html
5. Microsoft: Microsoft Accessibility http://www.microsoft.com/enable
6. W3C: Roadmap for Accessible Rich Internet Applications (Working Draft) (2006) http://www.w3.org/TR/aria-roadmap/

Making Multimedia Internet Content Accessible and Usable

Hisashi Miyashita, Hironobu Takagi, Daisuke Sato, and Chieko Asakawa

IBM Research, Tokyo Research Laboratory, Japan
{himi,takagih,daisuke,chie}@jp.ibm.com

Abstract. Although multimedia content containing streaming media is now widely used on the World Wide Web, there exist considerable difficulties for blind users to access such content, due to its dynamic changes, keyboard inoperability, and audio interference with the speech from assistive software. In particular, the third problem of audio interference is serious for blind users, since multimedia content often contains streaming media such as video and music which continuously play sounds, and thus they cannot hear the speech, which is masked by the loud media.

In this paper, we propose a new accessible browser that can directly manipulate such multimedia content. In order to control Flash contents, our browser relies on a transcoding HTTP proxy to inject special scripts into the Flash content and then manipulates the embedded streaming media and sound objects via the injected scripts. By using our browser, users can easily turn the volume up or down, play, stop, or pause the streaming media with shortcut keys. Since the users do not need to focus on buttons or sliders for these operations, they can immediately stop or fade out the intrusive media when listening to speech from assistive software.

1 Introduction

Since the Web is extending to a wider range of applications, in new Web sites, multimedia content is becoming increasingly popular for richer user experiences. Many prominent Web sites such as YouTube[1], Yahoo! Video[2], and MTV[3] distribute multimedia content and are accessed by enormous number of users. The multimedia content we discuss here has the following characteristics.

Dynamic changes. Most of the traditional HTML webpages are static. That is, the rendered information is not subject to change once the page is loaded. Unlike static contents, owing to the programmability of JavaScript or Flash with scripting, multimedia contents can change dynamically in response to various inputs.

Interactive user interface. Users want to interact with multimedia content for a variety of reasons, such as selecting movies, playing or stopping the media, and controlling the volume. Therefore, such multimedia content often has its own interactive user interfaces using a mouse or keyboard.

[1] http://www.youtube.com
[2] http://movies.yahoo.com
[3] http://www.mtv.com

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 98–107, 2007.
© Springer-Verlag Berlin Heidelberg 2007

Streaming media. Much multimedia content provides streaming media such as a movie or music in addition to text, GUI, or document information. Such streaming media is unique among media types in the sense that it delivers transient sounds and images in response to user's operations.

However, with typical accessible user agents such as screen readers and self-talking browsers, this multimedia content is hard for visually impaired people to access due to the following issues.

Issue 1: User agents miss dynamic changes. Most of the dynamic changes in multimedia content are simply ignored by many accessible user agents. For example, many media players show the current time of the movie or music but many accessible user agents cannot handle such dynamically changing information. This is partly because Web browsers or Flash players do not notify the assistive software components of the changes.

Issue 2: Content cannot be controlled using the keyboard. Much multimedia content is heavily dependent on mouse operations rather than keyboard operations. For example, volume sliders in a media player may not be controlled with the keyboard. Whether or not the content accepts keyboard operations often depends on scripts in the content and much multimedia content does not support keyboard operations.

Issue 3: Streaming media interferes with the synthesized speech Since streaming media emits sounds, it interferes with the voices synthesized by accessible user agents. Users cannot easily listen to mixed synthesized voices and media sounds. By controlling the volume of the streaming media, we can considerably reduce this problem. However, none of the user agents can directly control streaming media without using the appropriate button or slider.

These issues are critical for blind people accessing multimedia content on the Internet. According to the User Agent Accessibility Guidelines 1.0 [9], multimedia players shall allow users to start, stop, pause, and navigate multimedia and support full keyboard access. Thus, solving these problems is quite important in order to make multimedia content accessible. To the best of our knowledge, in the existing studies on the Web accessibility, although multimedia content has been discussed, this paper describes the first attempt to make the existing Internet multimedia content accessible.

In order to address these issues, we are now trying to make a multimedia accessible browser. In this paper, we mainly focus on Issues 2 and 3. For these issues, our browser can directly manipulate the media player for Flash content by using the technology discussed in [13]. Since our browser directly accesses the internal object model of Flash, we can control media players without any mouse operations. In addition, with our browser, users can immediately change the volume of streaming media at any time without focusing on a volume slider or a mute button. This does not affect the volume of the speech synthesized by the browser. Therefore, users can reduce the volume of a movie or music when they have problems hearing the synthesized speech. For Issue 1, although some activities are now being reported [14,1], since this issue involves many difficult problems involving detecting changes and properly notifying users, we are not addressing it in this paper.

The rest of this paper is organized as follows. In Section 2, we examine the problems that browsing typical multimedia content creates for a blind user by describing a

motivating example. Our solutions to these problems appear in Section 3, where we describe our multimedia accessible browser, its architecture, and the current implementation. In Sections 4 and 5, we discuss some related work and our conclusions.

2 Accessibility Issues of Internet Multimedia Contents

In this section, we consider a motivating example of actual multimedia content in order to clarify the problems.

We are considering a movie distribution site, which is one of the most popular web sites, delivering over 100 million video clips per day to users. In Figure 1, we show an example of the multimedia content this site delivers. In such content, a movie player is embedded in the upper-left corner, and the rest is used for showing DHTML content related to the selected movie. In order to control the movie, these contents provide some Flash widgets at the bottom of the media player. At the left side, we see a play/pause toggle button and skip buttons and at the right side, we see a volume slider and a mute button. In the center, we see the current time and the total time of the movie.

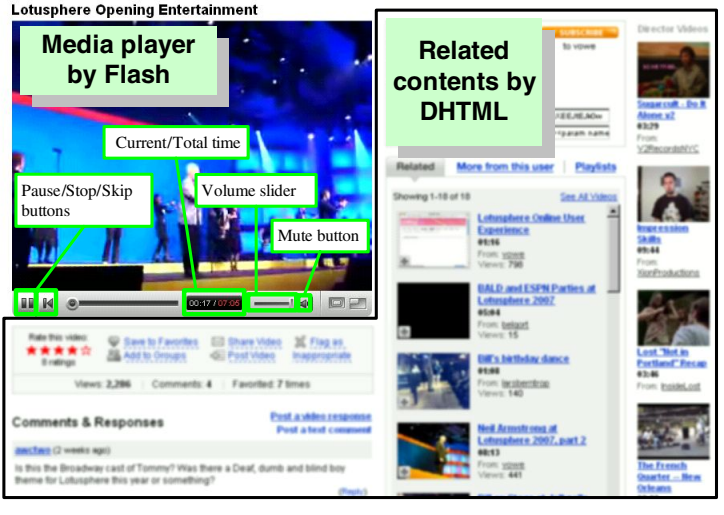


Fig. 1. An example of multimedia contents

For bind people using assistive talking systems such as screen readers and self-talking browsers, such content is difficult to access for the following four reasons.

- The movie starts immediately after the page is loaded (Issue 3). Thus, many blind users using assistive talking systems will get confused since the synthesized speech for guiding the users is suddenly interfered with by the audio of the movie.
- None of the widgets in this media player have any text information (Issue 2). Therefore, typical screen readers only say “Button one, Button two, ...” so that blind users cannot understand what button is for playing or stopping the movie.

– The volume control slider at the right side cannot be controlled with the keyboard (Issue 2) even if the user knew the meaning of the widget, because this player only accepts mouse operations for this slider. Therefore, blind users cannot change the volume of the movie.
– It needs considerable effort for a blind user to control the movie (Issue 3). In order to listen to synthesized speech to navigate in the DHTML part, users need to stop or mute the movie. However, to press the stop or mute button, they have to suspend the navigation in the DHTML part, select the appropriate button, and then push the enter key. These operations are quite irritating for the users.

Based on our investigation of the several sites hosting multimedia content, we found this site is not unusually inaccessible. Almost all of the sites have the similar problems. We believe this situation is partly because making Flash content accessible is not yet a priority [4] and sites that heavily rely on Flash tend to exploit the fancier GUI features, such as the streaming feature with FLV (FLash Video), and the eye-catching vector graphic animation features.

3 Multimedia Accessible Browser

As already shown, multimedia content on the Internet is highly unfriendly for blind users. Therefore, we are now making an accessible multimedia browser, which we call "aDesigner Runtime". Our browser can analyze multimedia content in Flash and directly control the streaming and sound objects. By taking this approach, users are not forced to manipulate the buttons and sliders offered by the multimedia content. Instead they can use predefined shortcut keys to control the media. In this section, we describe the detailed architecture of our browser.

3.1 Overall Architecture

The aDesigner Runtime is designed as a self-talking browser like Home Page Reader [3]. In Figure 2, we show the schematic diagram of our browser.

Our browser first injects special scripts into the incoming multimedia Flash content in a special transcoding HTTP proxy, using the same technology discussed in [13]. Then, the base browser runtime such as Microsoft Internet Explorer with the Adobe Flash Player receives the transcoded content and renders it. Our browser internally has a unified DOM[4]-based interface, which comes from the HTML DOM provided by Internet Explorer and the Flash objects provided by the injected scripts in the Flash contents. Since these DOM objects are integrated into one tree, any program using this interface can access the documents in a uniform way.

By using this DOM-based interface, the user interface part (UI part) provides a speech interface for users. From the UI part, users can navigate in the content and control the media player with the keyboard. Fixed keyboard shortcuts are specially assigned as media player controls, such as play, stop, pause, volume up, volume down, and mute the media. Thus, users can immediately manipulate the actively playing media.

[4] Document Object Model. See http://www.w3.org/DOM/ for details

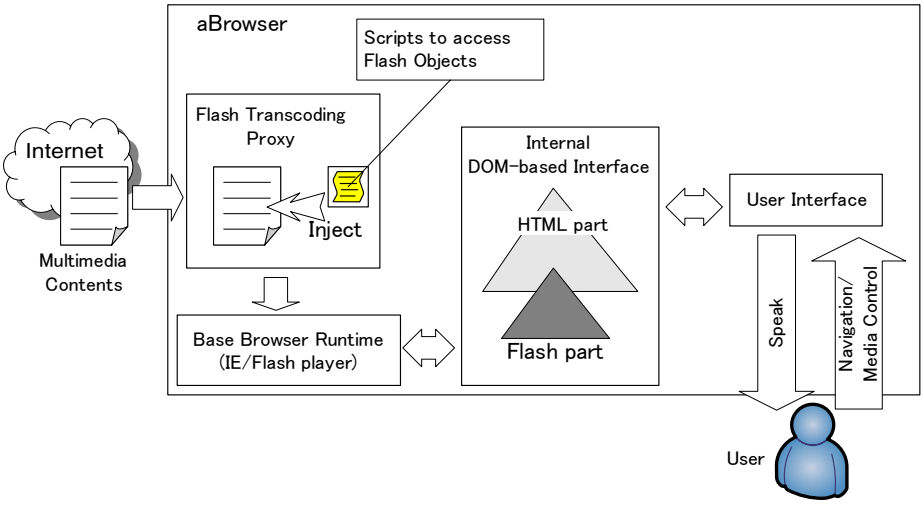


Fig. 2. Overall architecture

3.2 Flash Transcoding Proxy

Our browser uses a special transcoding HTTP Proxy to inject some ActionScript (AS) into the incoming Flash content. The injected AS has two roles: 1) it transforms the Flash objects into a DOM (e.g. a treelike object model), 2) it searches for streaming media objects (e.g. in the `NetStream` class) and sound objects (e.g. in the `Sound` class) and invokes some methods in these objects.

Since Role 1) is discussed in detail in [13], here we only add a few comments about it. The AS embedded in Flash contents traverses the structure of Flash objects from the root stage such as `_level0`. By limiting this traversal to visible objects, we can form a tree in which each node represents a visible Flash object. By using this tree, we can provide a DOM-base interface for Flash content.

For Role 2), we try to traverse all of the objects in the Flash content in order to find as many as possible of the streaming media objects and sound objects. Therefore, we traverse the `_global` property as well as the root stage. We also look into the nonvisual objects. Since the structure built of entire objects in Flash content may contain cyclic references, we track whether or not the next object has already been visited. This may still fail since objects may be stored in local variables where we cannot access from the external AS. However, in our experiments, we were able to all of the critical objects using these steps.

Once we find out the streaming media objects, we can invoke the `pause` method to play or pause the streaming media, and use the `seek` method to set the current position. By accessing the sound objects, we can use the `setVolume` method to set the volume of the media. Sometimes, Flash content contains more than one sound objects, not all of which are linked to streaming media objects. Even in such cases, we can adjust the volume of all of the sound objects with keeping the ratios among their volumes. This treatment is comparable to controlling a master volume control of all of the sound

sources in the Flash content and helps blind users listen to the synthesized speech from the assistive software, including aDesigner Runtime itself and other screen readers.

3.3 DOM-Based Interface

The aDesigner Runtime internally provides a DOM-based interface by accessing the HTML DOM supported by Internet Explorer and the DOM-based interface by the injected AS explained in Section 3.2. The reasons why we choose DOM as the common interface for the accessible browser are as follows.

- DOM is a well accepted standard for Web developers, since Web browsers and Web application frameworks support DOM for document manipulation.
- By leveraging the tree structure of a DOM, we can seamlessly integrate totally different content such as HTML and Flash, in a single DOM tree. Through this integration, the component using this interface (the UI part) does not need special considerations of their differences when we navigate in the DOM tree.
- By using the DOM, we can leave room for other Web-based technologies such as XPath and XSLT. For example, XPath can easily be introduced by using the DOM interface, which is greatly helpful for queries and for evaluating multimedia content.

This DOM-based interface provides some extensions as well as the standard DOM functions. These extensions can be classified into two groups: (a) navigation functions and (b) multimedia-content control functions. By using (a), we can click a node, obtain a heading node, and do some other tasks to provide an accessible UI. By using (b), we can search for streaming media objects and sound objects, and control the media by accessing them.

3.4 User Interface

The UI part of our browser is designed as a distinct component. This approach has the clear merit that we can present another UI by just replacing this component. For example, we could use a Braille display for blind users to navigate and control multimedia content or use tactile devices to quickly present lots of information. Otherwise, we would have to provide formatted text to other assistive programs such as screen readers by showing it in an appropriate window. This type of flexible user interface will be welcomed by the users who are familiar with their own favorite screen readers. The DOM-based interface is flexible enough for various UI components to provide good user experiences.

3.5 Implementation

We implemented aDesigner Runtime on top of the Eclipse Rich Client Platform (RCP)[5]. Most of the components in our browser are written in Java, though we use C++ to access the HTML DOM and the Adobe Flash Player via the COM (Component Object Model)

[5] http://wiki.eclipse.org/index.php/Rich_Client_Platform

and to access the Microsoft SAPI (Speech API)[6] for TTS (Text-to-Speech) feature. We also use ActionScript to write the scripts injected into the Flash content by the transcoding HTTP Proxy.

In Figure 3, we show a screen shot of our browser. In the current implementation, the user interface is based on tree navigation. That is, we can move up and down and go to the next and previous siblings in the tree made from the content. In the left panel, our browser shows the tree structure and the tree item currently selected. Every time the user moves in the tree, our browser announces the currently selected item. In addition, our browser has useful navigation commands such as "skip to heading" or "skip to next form input item," similar to typical accessible browsers such as Home Page Reader. Note that all of these features can be implemented through the DOM-base interface.

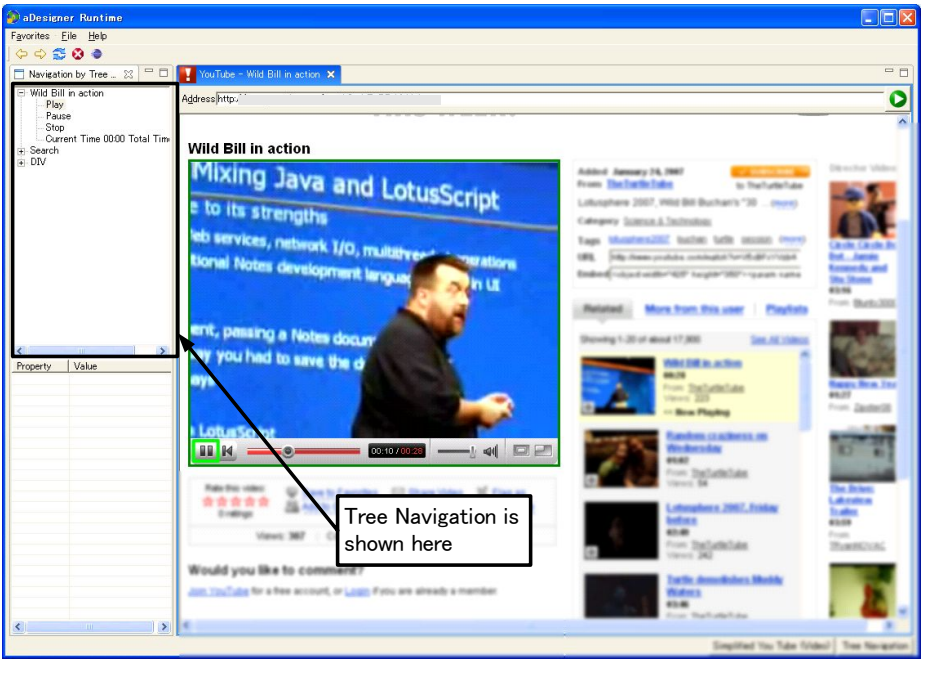


Fig. 3. A screenshot of aDesigner Runtime

Our browser is different from the others in that users can directly control embedded media by simply pressing shortcut keys. For example, by pressing `Ctrl+J`, `Ctrl+K`, or `Ctrl+M`, users can increase, decrease, or mute the volumes of the playing media, respectively. Likewise, by pressing `Ctrl+P`, `Ctrl+S`, or `Pause`, they can play, stop, or pause the media, respectively. Currently, our browser supports this direct media control for Flash and FLV contents, but we will also be able to support Windows Media Player by using the standard API, which will be easier than supporting Flash content.

[6] http://www.microsoft.com/speech/default.mspx

4 Related Work

4.1 Screen Readers and Self-talking Browsers

There exist a number of studies to access Web contents with special browsers for blind people [3,17,10]. However, there is relatively little research on making multimedia contents accessible. Recently, some accessible browsers and screen readers have begun to support multimedia content. For example, Home Page Reader [3] by IBM now supports multimedia applications such as Flash and Windows Media Player. Some screen readers such as JAWS[7] and Window-Eyes[8] also support such multimedia contents. However, all the existing softwares uses MSAA (Microsoft Active Accessibility)[9] to access Flash contents. Therefore, such software can only notify users using functions that are supported for multimedia content in MSAA. Thus, they do not have special functions to control multimedia contents as our browser does. It means that with these programs, we have to focus on the buttons or sliders that are shown by the multimedia content and use them to control the playing media.

Other interesting research coping with simultaneously presenting multimedia content and TTS is Streaming Speech[3] [7]. They proposed an extension to SMIL [2], a standardized multimedia integration language, in order to support 3D spatialization of multiple audio sources. Their approach makes use of the human ability to select multiple audio sources by 3D placement, in contrast to our approach. However, it has not yet been proved whether or not 3D spatialization of multimedia content is accessible for blind users.

4.2 Web Content Adaptation

Adaptation of Web content by transforming or analyzing it is another approach to provide more comprehensible information to a wide range of users [12,15,11,8,6,16]. Also using this approach, there are a few research projects on multimedia content. For example, Rousseau *et al.* proposed an adaptable multimedia presentation framework [12] by extending HTML. Thus, multimedia content has already been described using an extended format. Although HearSay [11] automatically analyzes Web pages and enables audio navigation, it does not address multimedia content. DeMeglio *et al.* showed a customizable multimedia player, called *AMIS*, for DAISY content [6], which has to be natively accessible. Weimann *et al.* proposed a time-dependent presentation system, called *MultiReader*, for multimedia content. MultiReader uses XSL stylesheets to reorganize content into accessible forms and also offers multimedia synchronization methods, which can be specified in JavaScript. However, this system also requires a special transcoding rule for each piece of content.

5 Conclusions and Future Work

We described a new accessible browser that tackles some of the issues when browsing multimedia content. Although lots of multimedia content poses problems for blind users

[7] http://www.freedomscientific.com/fs_products/software_jaws.asp
[8] http://www.gwmicro.com/Window-Eyes/
[9] http://msdn.microsoft.com/library/default.asp?url=/library/en-us/msaa/msaastart_9w2t.asp

using assistive speech software, our browser is an important step toward a multimedia-accessible browser. Our browser can directly control multimedia content without considering any special context such as finding special buttons and pressing them. In particular, controlling the volume of multimedia content is indispensable for blind users who need to hear the speech from their assistive software. Our browser can directly control multimedia Flash content by injecting AS code into that Flash content with a transcoding HTTP proxy. With our browser, users can easily do basic operations on the media such as changing the volume with unchanging shortcut keys.

As we have discussed so far, there is a long way to go to make multimedia content accessible. One of the important issues is about more flexible media control of multimedia content. Currently, our browser meets a minimum requirement of multimedia control such as play, stop, pause and volume controls. Since many blind users have prominent ability to recognize high-speed audio [5], in order to leverage their ability for more efficient browsing, we believe that more flexible user interface of multimedia content such as variable speed control and audio skimming control which extracts important sections is valuable especially for blind users to effectively access such content.

The other significant issue we should cope with is dynamic changes of multimedia content. Lots of multimedia content heavily uses dynamic visual changes. Since most of those changes are only eye candy, notifying visually impaired people of all of the changes is simply annoying and unhelpful. Some activities such as WAI-ARIA [14] are now tackling this problem by defining metadata for DHTML (Dynamic HTML). With WAI-ARIA, Web developers can add some metadata to DHTML content to articulate changes. However, since WAI-ARIA is defined as a set of internal metadata specific to DHTML, it has two limitations. First, we cannot apply it to other media types than DHTML, such as Flash. Second, we have to change the content itself to provide the metadata.

Therefore, we are now considering external metadata which can be applied to various types of media including Flash and DHTML. We do not have to change the multimedia content itself to attach such external metadata. With such metadata, accessible browsers can appropriately notify the users about dynamic changes of the contents.

References

1. Accessibility/IAccessible2 http://www.linux-foundation.org/en/Accessibility/IAccessible2
2. Synchronized multimedia integration language (SMIL 2.0) - [2nd edn]. W3C Recommendation (January 07, 2005) http://www.w3.org/TR/2005/REC-SMIL2-20050107/
3. Asakawa, C., Itoh, T.: User interface of a home page reader. In: Assets '98: Proceedings of the third international ACM conference on Assistive technologies, pp. 149–156. ACM Press, New York (1998)
4. Asakawa, C., Itoh, T., Takagi, H., Miyashita, H. Accessibility evaluation for multimedia content. In: UAHCI (2007)
5. Asakawa, C., Takagi, H., Ino, S., Ifukube, T.: Maximum listening speeds for the blind. In: International Community for Auditory Display 2003, pp. 276–279 (2003)
6. DeMeglio, M., Häkkinen, M. T., Kawamura, H.: Accessible interface design: Adaptive multimedia information system (amis). In: ICCHP 2002, pp. 406–412 (2002)

7. Goose, S., Kodlahalli, S., Pechter, W., Hjelsvold, R.: Streaming speech[3]: a framework for generating and streaming 3D text-to-speech and audio presentations to wireless PDAs as specified using extensions to SMIL. In: WWW pp. 37–44 (2002)
8. Huang, A.W., Sundaresan, N.: Aurora: a conceptual model for web-content adaptation to support the universal usability of web-based services. In: CUU '00: Proceedings on the 2000 conference on Universal Usability, pp. 124–131. ACM Press, New York, USA (2000)
9. Jacobs, I., Gunderson, J., Hansen, E.: User agent accessibility guidelines 1.0. W3C Recommendation (December 17, 2002) http://www.w3.org/TR/UAAG10/
10. King, A., Evans, G., Webbie B.P: a web browser for visually impaired people. In: Proceedings of the 2nd Cambridge Workshop on Universal Access and Assistive Technology, pp. 35–44 (2004)
11. Ramakrishnan, I.V., Stent, A., Yang, G.: HearSay: enabling audio browsing on hypertext content. In: WWW '04: Proceedings of the 13th international conference on World Wide Web, pp. 80–89. ACM Press, New York, USA (2004)
12. Rousseau, F., García-Macías, J.A., de Lima, J.V., Duda, A.: User adaptable multimedia presentations for the world wide web. Computer Networks 31, 11–16, 1273–1290 (1999)
13. Saito, S., Takagi, H., Asakawa, C.: Transforming flash to xml for accessibility evaluations. In: ASSETS, pp. 157–164 (2006)
14. Schwerdtfeger, R.: Roadmap for accessible rich internet applications (WAI-ARIA roadmap). W3C Working Draft (December 20, 2006) http://www.w3.org/TR/aria-roadmap/
15. Takagi, H., Asakawa, C., Fukuda, K., Maeda, J.: Site-wide annotation: reconstructing existing pages to be accessible. In: ASSETS, pp. 81–88 (2002)
16. Weimann, K., Langer, I., Weber, G.: Adaptation of multimedia browsing techniques. In: ICCHP, pp. 135–142 (2004)
17. Zajicek, M., Powell, C., Reeves, C.: A web navigation tool for the blind. In: Assets '98: Proceedings of the third international ACM conference on Assistive technologies, pp. 204–206. ACM Press, New York, USA (1998)

An Evaluation of Mobile 3D-Based Interaction with Complex Multimedia Environments

Ali A. Nazari Shirehjini and Jafar Movahedi

Fraunhofer-Institute for Computer Graphics,
Darmstadt, Germany
ali.nazari@igd.fraunhofer.de
http://www.igd.fhg.de/igd-a1/

Abstract. This paper describes the evaluation of an intuitive mobile environment controller. Through its 3D-based user interface, the PDA based controller provides the user with mobile access to her/his physical environment and ambient media. The goal of this usability evaluation is to test the performance and analyse the user acceptance of 3D-based interaction systems compared with traditional WIMP-based controllers. At first this paper describes the interaction model which has been evaluated. Then it explains the experiment and the approach of the evaluation. Finally we present and discuss our results.

1 A Novel 3D-Based Interaction Model for Human-Environment-Interaction

Ambient Intelligence (AmI) is the vision of a world where we are surrounded by a huge amount of intelligent and small things, which are seamlessly embedded in the user‘s environment [1]. They pro-actively support the user‘s everyday activities. In this world one major question is how to interact with all those devices? How to access physical devices in an unfamiliar environment without having knowledge about the technical infrastructure such as device addresses and IP-numbers?

Existing interaction systems for multimedia environments (like meeting or lecture rooms) are mostly based on complex menus and icon-based user interfaces (e.g., see fig. 1). In order to access a device, the user has to browse complex menu hierarchies and select a device based on its unique name or device address. Especially in unfamiliar environments, this is a challenging task (due to cognitive overload).

The Personal Environment Controller (PECo) provides a novel interaction metaphor to control intelligent environments. Its goal is to allow the user to easily access virtual and physical objects surrounding him. PECo uses an automatically created 3D visualization of the environment. Entering a room PECo discovers the infrastructure as well as available devices and constructs the integrated user interface. The user can easily identify a device within his environment based on its position, orientation, and form, and directly manipulate it within the 3D-scene. For example, he can simply click on a 3D object to turn on a light. The 3D-based user interface allows the user to access the infrastructure without having knowledge about specific device names,

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 108–115, 2007.
© Springer-Verlag Berlin Heidelberg 2007

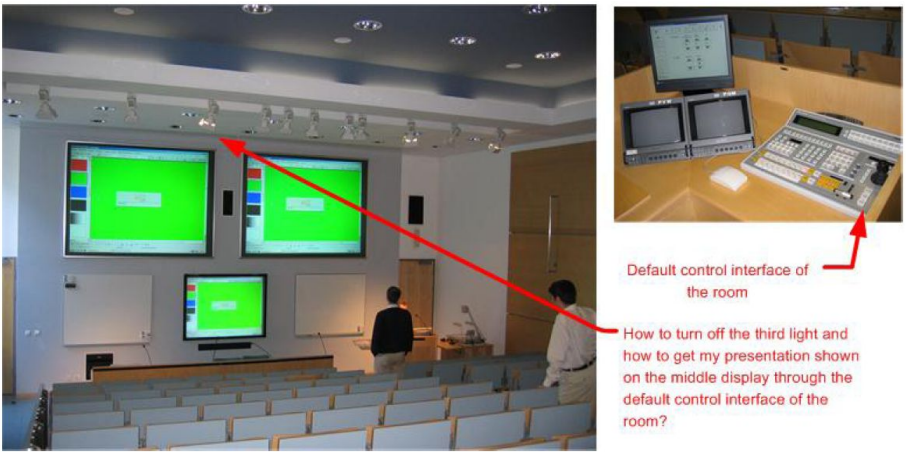


Fig. 1. Challenges of Human-Environment-Interaction and the complex structure of existing control systems (the control panel right)

IP-Numbers, URLs, etc. Figure 2 shows a complex meeting room and the corresponding user interface within the PECo system.

Furthermore PECo amalgamates the user's virtual media repository and the physical environment into a personal environment. By interconnecting these two worlds, the user can drag a presentation from her/his virtual media repository and drop it onto a representation of a real-world display with just one interaction step (drag&drop).

The interaction model has been discussed in detail in [2]. Some other aspects of the PECo user interface, the interaction concept, as well as the way how the 3D model is created dynamically for a new environment, are described in detail in [2, 3, 4, and 5].

2 The Evaluation

The evaluation of PECo has been done by a quantitative usability test and a complementary qualitative analysis of the software ergonomics based on ISO 9241/10 questionnaires [6].

Within the scope of a quantitative usability test, the task completion time and the rate of errors have been analysed. For this reason the subjects have used two interaction systems to perform activities within a meeting room. This allowed us to compare the usability of the mobile 3D-based PECo system against a stationary central control panel (CCP) which is based on traditional menus and icon-based metaphors.

The user evaluation has been performed in a traditional meeting room (see fig. 1). The usability test analysed most important activities of a presentation scenario. We selected the test situations and the activities involved with them from a situation-concept and a hierarchical task model for presentation scenarios see [7–9]).

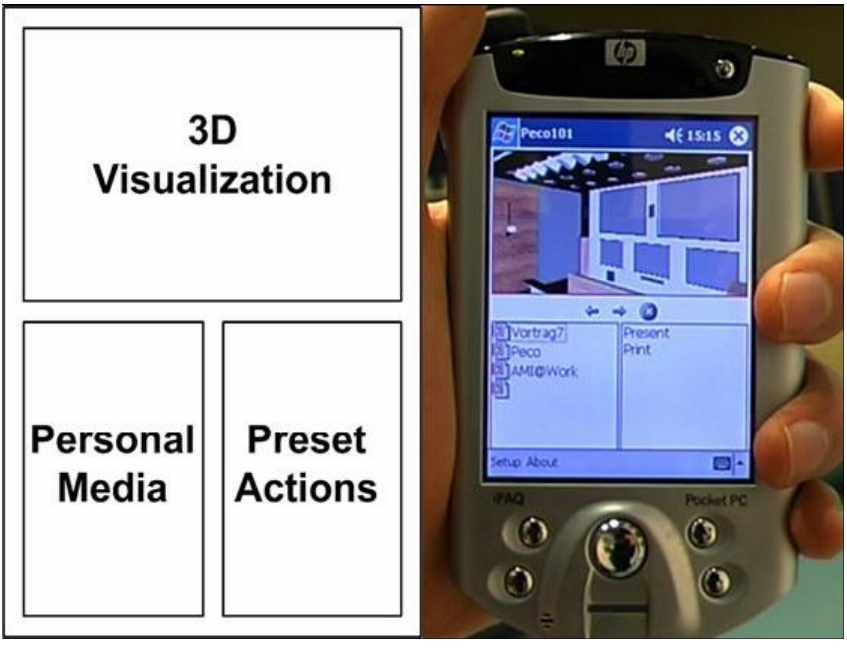


Fig. 2. The PECo-system uses a 3D-based interface to provide access to complex multimedia environments

A total of 46 subjects performed 41 different test cases (e.g., "set up the room for a presentation!" or "present your thesis-slides on the back-projection system!"); all cases included several activities within three different situations (Presentation, Setup, Explanation). Each activity has been performed three times with the PECo system and three times with the central control panel. By doing so we were able to analyse the learnability / the learning curve for each of the interaction system. Some activities have also been performed using manual controls (e.g. light switch). Before performing an activity, we prepared the room and re-started the interaction systems. Each subject has been introduced to both of the control systems. Two scientists observed them during their tasks. One of them measured task completion time. The other one counted the subject's error rate. We avoided having several test users in the room at the same time.

After performing the quantitative usability test (task completion time and error rate) we interviewed 38 of the subjects using the ISO-questionnaires to analyze the suitability of PECo, the self-descriptiveness of the user interface, controllability, conformity with user expectations, error tolerance, suitability for individualization and suitability for learning. The structure of those questionnaires is described in [6]. The interviews have been used to perform a qualitative analysis about the suitability of PECo. Minor modifications have been done in order to adopt the questionnaires to our system.

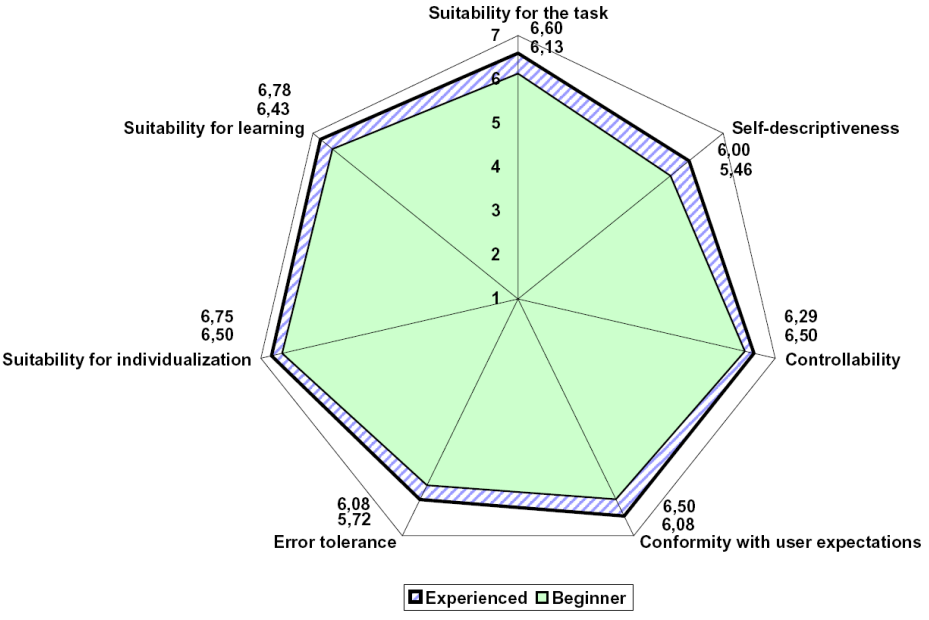


Fig. 3. Results of the ISO 9241/10 questionnaire showing the user interface "quality" of PECo. We show the results for two different user groups; a 3D-user interface beginners group and an experienced 3D-interface users group. All users have already worked before with PDA and touch-screens.

3 Results and Discussion

We present a study that evaluates conventional control systems and 3D- based user interface for accessing physical environments within presentation scenarios.

We are presenting results from two studies. The first study precisely demonstrates the performance of each interaction system for a given task. The second study shows the general usability aspects of the PECo system.

Our results show that the 3D-based interaction was significantly faster and required less mental effort than conventional interaction systems; this was true for all situations and for all performed activities. For example, the subjects performed the tasks for environment management – like light or projector control – just by clicking once into the 3D-scene. This step took 2.15 sec in average.

In order to perform the same activity with the CCP they needed more then 10 sec. We also analysed the learning curve for each interaction systems. The learnability for PECo is very good in comparison to the other systems.

Figure 4 shows the task completion time for 4 activities which have been performed under the situations "Setup" and "Presentation". For the required Human-Environment-Interaction manual control systems like switches, the central control panel (CCP), and the PECo system have been used. Figure 5 shows the task completion time for 8 different activities which have been performed using the different

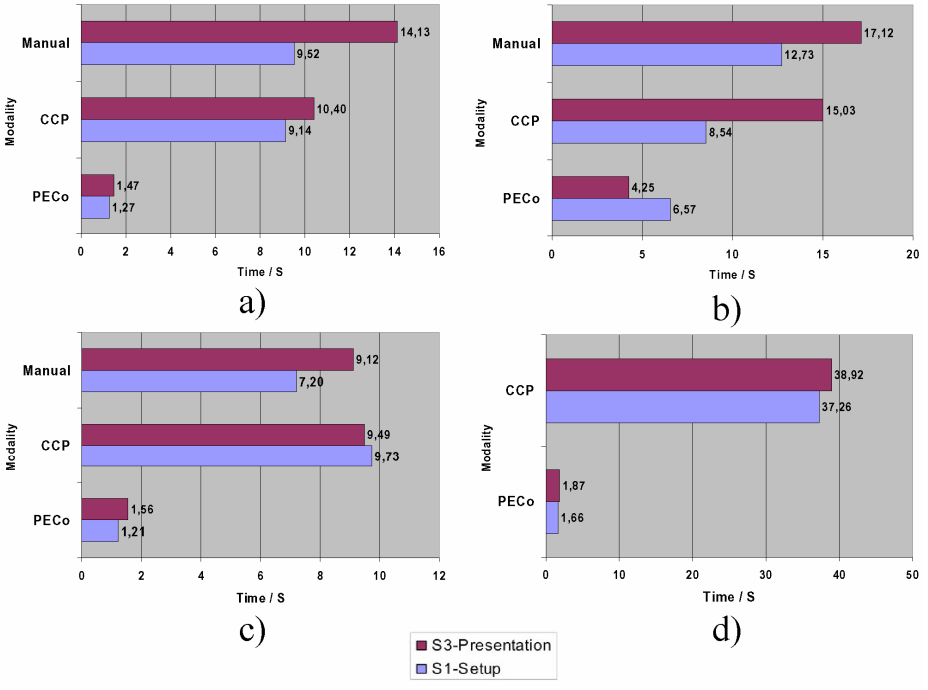


Fig. 4. The picture comprises the task completion time within the situations "Setup" and "Presentation" for 4 activities: a) Audience Light control, b) blind shutter control, c) Panel Light control, d) presentation control

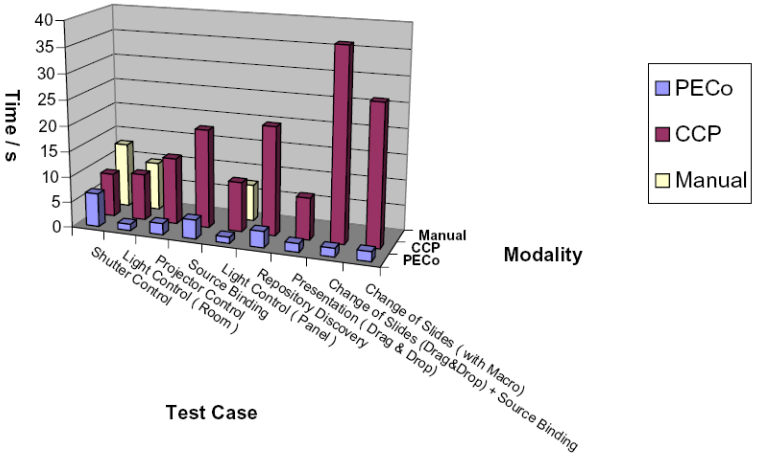
control systems under the "Setup" situation. The activity "presentation change" was performed using two different metaphors of the PECo system. These results show that the overall performance of the analyzed activities and situations is significantly higher for PECo compared to CCP or manual control.

Additionally the results of the questionnaires show (see fig. 3) that the interaction design of the PECo is easy to understand and easy to use.

Moreover, the relationship between user experience and efficiency can be expressed by learning curve. As individuals get more experienced at a task, they usually become more efficient at them. The learning curve states that the more times a task have been performed, the less time will be required on each subsequent iteration.

Figure 6 shows, the learning curve of performing four tasks using two interaction systems, PECo, and CCP. The user performed three times the specific activity (1^{st} – 3^{rd} iteration) within the "setup" situation. The rest of the test case (iterations 4^{th} – 6^{th}) have been performed and measured under the "presentation" situation. As you can see the required time for performing the test number 4 increases in all diagrams, since there is a time gap (~ 30 minutes) between running test 3 and 4.

Figure 6.a, shows the learning curve of presentation rendering task. The average time of performing this task in test2 is rapidly decreased. Actually, the "presentation rendering" task is a set of seven tasks; projection selection, projector on/off, source



	Shutter Control	Light Control (Room)	Projector Control	Source Binding	Light Control (Panel)	Repository Discovery	Presentation (Drag & Drop)	Change of Slides (Drag&Drop)	Change of Slides (with Macro)
PECo	6,57	1,27	2,24	3,74	1,21	3,19	1,70	1,66	1,92
CCP	8,54	9,14	13,02	19,29	9,73	21,20	8,20	37,26	27,50
Manual	12,73	9,52			7,20				

Fig. 5. This diagram shows the task completion time for the "Setup" situation whereby PECo-based control is compared to CCP and manual control

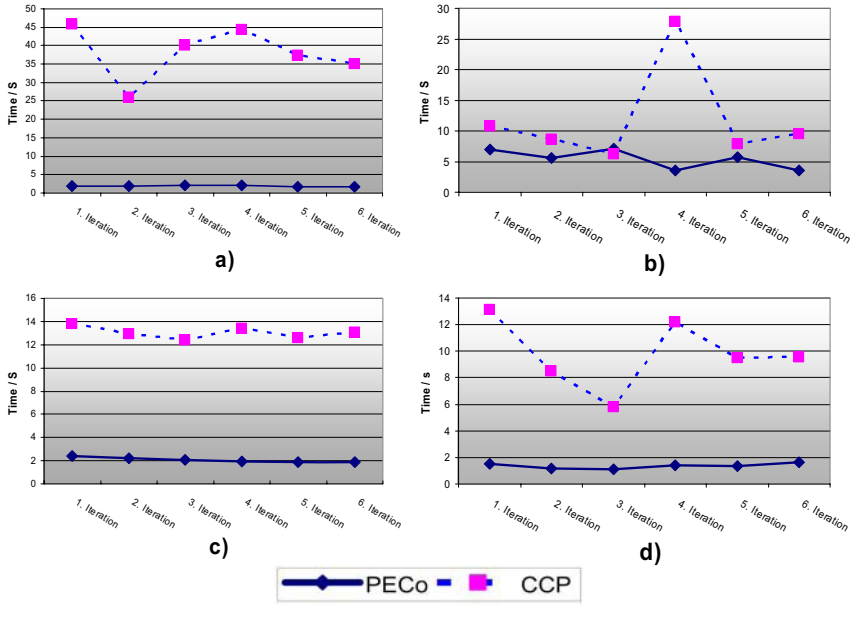


Fig. 6. The learning curve of the PECo interaction system compared to the CCP interaction systems for performing four typical tasks; a) Presentation rendering, b) Shutter movement, c) Projector control, and d) Audience light adjustment

browsing & selection, source binding, media browsing & selection, media opening, and media rendering. In test1/ test3, the user has been asked to choose notebook/camera as source and show its output on the display. It was very complicated and time consuming task to find notebook/camera source option within the CCP interaction system. But in test2, the users have asked to choose "Room PC" as source which was an easy task using the CCP interaction system. Hence required time for performing "presentation rendering" task making use of CCP in test 2 was very low, since most users were able to find the "Room PC" quite easily.

Figure 6.b, belongs to "shutter movement" task. The average time of performing this task for the test 4 is too much (~ 27 seconds). The reason for that lies in the complicated shutter control of the CCP interaction system. Some users have found the shutter interaction menu too late, while some users could not find the shutter interaction menu on the CCP at all.

The percentage of learning could also be calculated by the learning percentage formula (1) [10],

$$Y_x = Kx^{\log_2 b} \tag{1}$$

where K is the time of performing the desired task for the first test, Y_x is the time of performing the desired task for the x-th test, x is the test number, b is the learning percentage.

Making use of this formula, we have realized that the learning percentage of all aforementioned tasks using PECo is higher compared to CCP. As an example, the learning percentage (b) for performing projector control task using PECo system is 9% whereas it is just 2% for the CCP system.

Future research will focus on providing new metaphors and icons as well as improving the user interface look&feel. An extended usability evaluation will be conducted to investigate and compare different metaphors to perform a certain interaction step.

References

1. Kirste, T.: Ambient Intelligence and Appliance Ensembles: Architectural Considerations. In: Conference on Human Factors in Computing Systems (CHI), Netherlands (2004)
2. Shirehjini, A.A.N.: A Novel Interaction Metaphor for Personal Environment Control: Direct Manipulation of Physical Environment based on 3D Visualization. In: Computers & Graphics, Special Issue on Pervasive Computing and Ambient Intelligence, vol. 28, pp. 667–675. Elsevier, Amsterdam (2004)
3. Shirehjini, A.A.N.: A Generic UPnP Architecture for Ambient Intelligence Meeting Rooms and a Control Point Allowing for Integrated 2D and 3D Interaction. In: sOc-EUSAI '05: Proceedings of the 2005 joint conference on Smart objects and ambient intelligenc, pp. 207–212. ACM Press, New York (2005)
4. Shirehjini, A.A.N.: PECo: 3D-based Interaction with a UPnP Meeting Room, Austria, OCG (2005)
5. Shirehjini, A.A.N., Hellenschmidt, M., Kirste, T.: An integrated user interface providing unified access to intelligent environments and personal media. In: Proceedings of the 2nd European Union symposium on ambient intelligence, pp. 65–68. ACM Press, New York (2004)

6. Prümper, J., Anft, M.: Beurteilung von Software auf Grundlage der Internationalen Ergonomie-Norm ISO 9241/10. web (N.N.) http://www.bao.de
7. Shirehjini, A.A.N.: Situation Modelling: A Domain Analysis and User Study. In: IE '06: accepted for publication in the Proceedings of The 2nd IEE International Conference on Intelligent Environments (IE '06) (2006)
8. Shirehjini, A.A.N., Movahedi, J.: Situation modelling for ambient intelligence: Towards scenario-based methodology and formal representation of context. In: CSICC' 06: Proceedings of the 11th International CSI Computer Conference (2006)
9. Shirehjini, A.A.N.: Context-Awareness: DynAMITE Environment Model. In: ASWN2006: Accepted for publication in Proceedings of the 6th International Workshop on Applications and Services in Wireless Networks (2006)
10. Chase, R.B.: Operations management for competitive advantage, 9th edn. International edition: McGraw Hill/ Irwin (2001) ISBN 0-07-118030-3

Detection of Layout-Purpose TABLE Tags Based on Machine Learning

Hidehiko Okada and Taiki Miura

Kyoto Sangyo University
Kamigamo Motoyama, Kita-ku, Kyoto 603-8555, Japan
hidehiko@ics.kyoto-su.ac.jp

Abstract. To make webpages more accessible to people with disabilities, <table> tags should not be used as a means to layout document content. Therefore, to evaluate the accessibility of webpages, it should be checked whether the pages include layout-purpose <table> tags. Automated precise detection of layout-purpose <table> tags in HTML sources is still a research challenge because it requires further than simply checking whether specific tags and/or attributes of the tags are included in the sources. We propose a method for the detection that is based on machine learning. The proposed method derives a <table> tag classifier that deduces the purpose of a <table> tag: the classifier deduces whether a <table> tag is a layout-purpose one or a table-purpose one. We have developed a system that derives classification rules by ID3. The system derives a decision tree from a set of learning data (<table> tags of which the purposes are known) and classifies <table> tags in webpages under evaluation. Classification accuracy was evaluated by cross validation with 200 test data collected from the Web. Result of the evaluation revealed that 1) the tags can be roughly classified with attribute values of border, number of rows, number of tags that appear ahead of the <table> tag, and the nest of <table> tags (i.e., these attributes are more likely to appear in upper layers in decision trees), and 2) the accuracy rates are about 90% for the 200 test data.

Keywords: Web accessibility, automated checking, <table> tags, machine learning, ID3.

1 Introduction

To make webpages more accessible to people with disabilities, <table> tags should not be used as a means to visually layout document content: layouting contents by <table> tags may present problems when rendering to non-visual media [1],[2]. It is reported that the number of tables on pages doubled from 7 in 2000 to 14 in 2003, with most tables being used to control page layouts [3]. Therefore, to evaluate the accessibility of webpages, it should be checked whether the pages include layout-purpose <table> tags. Several methods and tools have been proposed and developed for web accessibility [4]-[14]. Accessibility tools are listed in [13],[14] and compared in [7]. Some tools detect deeply nested <table> tags that are likely to be layout-purpose ones. Still, a method for automated detection of layout-purpose <table> tags

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 116–123, 2007.
© Springer-Verlag Berlin Heidelberg 2007

in HTML sources is a challenge: it requires further than simply checking whether specific tags and/or attributes of the tags are included in the sources.

We propose a method for the detection that is based on machine learning. The proposed method derives a <table> tag classifier that classifies the purpose of the tag: the classifier deduces whether a <table> tag is a layout-purpose one or a table-purpose one. We describe our system that implements the proposed method and report a result of experiment for evaluating classification accuracy.

2 Proposed Method

2.1 Basic Idea

<Table> tag is used for (1) expressing data tables as the tags are originally designed for or (2) adjusting the layout of document contents. In the case where a <table> tag is used for the table-purpose, the data in the same row or column are semantically related with each other (e.g., values of the same property for several data items, or values of several properties for the same data item) and the relation can be expressed by row/column headers. On the other hand, in the case where a <table> tag is used for the layout-purpose, the data in the same row or column may not be semantically related. Thus, to deduce the purpose of a <table> tag in a fundamental approach, it should be analyzed whether or not the data in the same row/column of the table are the semantically related ones. To make the analyses automated by a computer program requires a method for semantic analyses of table contents, but the automated semantic analyses with enough precision and independency for any kinds of webpages is hard to achieve.

Our basic idea focuses on machine-readable design attributes of a table instead of analyzing semantics of data in a table. If some common design pattern is found in some attribute values of layout-purpose <table> tags and another common design pattern is found in the attribute values of in table-purpose <table> tags, the two purposes can be discriminated by denoting classification rules based on the design patterns. However, it is unknown whether we can find such design patterns, and even if we can, it will be hard for us to manually investigate common patterns by analyzing design attribute values in a large number of <table> tag instances and denote sufficient rules to precisely classify the tags. In our research, we manually investigate design attributes only: to automatically derive classification rules, we utilize a machine learning method.

Among several kinds of machine learning methods available, we here utilize ID3 [15], a method for deriving a decision tree from a set of data instances. An advantage of a decision tree as a classifier over other forms of classifiers (e.g., a multi-layered neural net) is that classification rules are obtained as tree-formed explicit if-then rules and thus easy to read for human users of the method.

2.2 <Table> Tag Design Attributes for Classification Rules

We first collected and investigated <table> tag instances (webpage HTML sources including <table> tags) and extracted design attributes of <table> tag that seem to contribute to the classification. Webpages were collected from various categories in

Yahoo webpage directory so that the pages were not biased from the viewpoint of page categories. We manually judged the purpose of <table> tag instances in the collected pages. By this way, we collected 200 <table> tags a half of which were layout-purpose ones and the others table-purpose ones. We then extracted common design patterns for each set of <table> tags. The findings were as follows.

- Common design patterns in layout-purpose <table> tags
  - <Table> tags are nested.
  - Some cell(s) in the table include image(s).
  - The number of HTML tags that appear before the <table> tag in the source is small.
  - Some cells in the table are spanned.
  - The width and/or height are/is specified.
- Common design patterns in table-purpose <table> tags
  - The table has visible border lines.
  - The table includes many rows.
  - <Th> tags for row/column headers are included.
  - Table titles are specified.

By denoting classification rules based on these common design patterns, it will be possible to deduce the purposes of <table> tags to a certain extent. To derive the rules in a form of decision tree by ID3, we determined 10 attributes in Table 1.

Table 1. Attributes for Decision Tree

Name	Meaning	Values
border	Whether the table has visible border lines.	Binary
caption	Whether the table has a caption.	
height	Whether the table height is specified.	
img	Whether a cell in the table includes an item of image data.	
nest	Whether the table includes a nested table in itself.	
span	Whether some cells are spanned.	
th	Whether a row/column has a title header for the data in the row/column.	
width	Whether the table width is specified.	
num_tag	The number of HTML tags that appear before the <table> tags in the source.	Positive integer
num_tr	The number of rows in the table.	

- Binary values for the eight attributes of “border”-“width” mean whether or not the table has visible borders, captions, etc. For example, if the border attribute is specified for the <table> tag and the attribute value is 1 or more then border = Y, else border = N.
- A value of the attribute “num_tag” is the number of HTML tags that appear before the <table> tags in the source. In counting the tags, those in the header part are not counted.
- A value of the attribute “num_tr” is the number of rows (<tr> tags) in the table.

2.3 System Configuration for the Proposed Method

Fig. 1 shows the system configuration for our method. In the learning phase, the method derives rules for classifying <table> tags from a set of webpages in which layout/table-purpose <table> tags are included. First, for each <table> tag in the learning data, attribute values are obtained by analyzing the webpage HTML source in which the tag is included. The purpose of each tag in the learning data is given by manual judgments in collecting the learning data. From these data of attribute values and given purposes, classification rules are derived by a machine learning method. In the case where ID3 is applied as the learning method, the rules are obtained as a decision tree (i.e., tree-formed if-then rules). In the checking phase, <table> tags (of which the purpose is unknown) are classified by the rules obtained in the learning phase. Attribute values of each <table> tag for the classification are obtained by the same way as in the learning phase (the <table> attribute analyses module in Fig. 1 is shared by the learning and checking subsystems). The obtained attribute values are applied to if parts of the rules, and the purpose of each <table> is deduced by the rule of which the if part is satisfied.

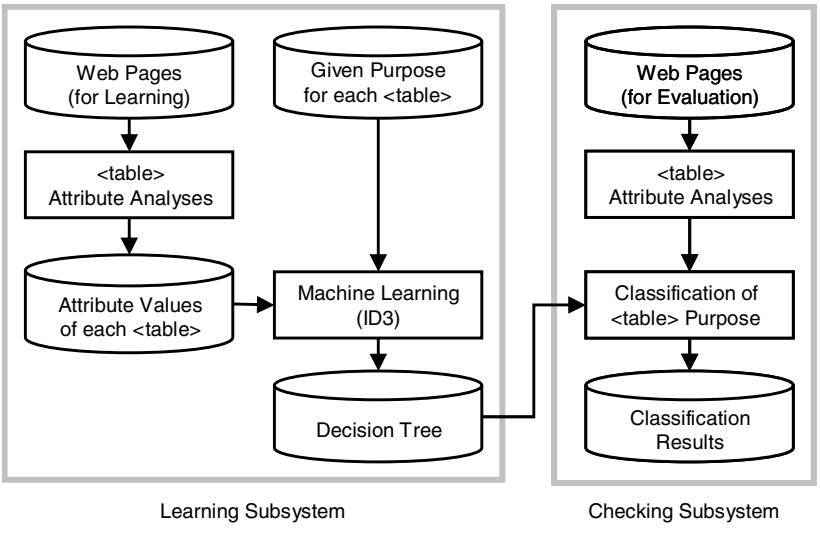


Fig. 1. System Configuration for the Proposed Method

3 Evaluation of the Proposed Method

3.1 Evaluation Method

To evaluate the effectiveness of our method, we investigate the classification accuracy by 10-fold cross validation (CV) with the 200 <table> tags we collected (see subsection 2.2). The 200 tags are randomly divided into 10 groups (G_1, G_2, …, G_{10}).

A decision tree is derived by ID3 applying to 180 <table> tags in 9 groups excluding G_i, and the classification accuracy is checked with 20 <table> tags in G_i (i=1,2,…,10). The accuracy rate is calculated as follows.

$$\text{Accuracy Rate} = (\text{Number of <table> tags correctly classified})/20 \quad (1)$$

Ten values of the accuracy rate are obtained by a trial of 10-fold CV. We test the 10-fold CV three times and statistically investigate the accuracy rates.

3.2 Decision Trees Obtained by ID3

By the three trials of 10-fold CVs, 30 decision trees are obtained in total. Examples of the decision trees are shown in Tables 2-4 (only the nodes in the depth ≤ 3 are shown). Values in the "No." column are the serial numbers of the nodes where the #0 node is the root node. Tables 2-4 denote parent-child node relationships by indents in the "Rule Element" column. For example, in Table 2, the #1 and #8 nodes are child nodes of the #0 node and the #2 and #5 nodes are child nodes of the #1 node. Values in the "Rule Element" column are the conditions in if-parts of rules. Conditions that appear in a path from the root node to a leaf node are connected with AND. Values in the "Table" and "Layout" columns are the number of table/layout-purpose <table> tags in the learning data included in the node. For example, Table 2 shows the followings.

- The root node includes 89 table-purpose tags and 91 layout-purpose tags (see #0 node).
- Of the 180 <table> tags in the root node,
  - those with border=N are 86. Nine tags are table-purpose ones and the other 77 tags are layout-purpose ones (see #1 node), and
  - those with border=Y are 94. 80 tags are table-purpose ones and the other 14 tags are layout-purpose ones (see #8 node).

Table 2. Example (1) of Decision Trees Obtained in the Experiment

No.	Rule Element	Table	Layout	Total
0	(root)	89	91	180
1	border = N	9	77	86
2	num_tr ≤ 7	1	71	72
3	num_tag ≤ 12	0	66	66
4	num_tag > 12	1	5	6
5	num_tr > 7	8	6	14
6	nest = N	8	1	9
7	nest = Y	0	5	5
8	border = Y	80	14	94
9	nest = N	78	5	83
10	img = N	59	2	61
11	img = Y	19	3	22
12	nest = Y	2	9	11
13	height = N	0	5	5
14	height = Y	2	4	6

- Nodes of which the values in the "Table" column or the "Layout" column are leaf nodes. A leaf node corresponds to an if-then rule. For example, the #3 node denotes that all tags that meet the condition:
  - (border = N) and (num_tr ≤ 7) and (num_tag ≤ 12)
  - are layout-purpose ones. Thus, the following rule is obtained from the #3 node.
  - If (border = N) and (num_tr ≤ 7) and (num_tag ≤ 12) then the tag is a layout-purpose one.

Table 3. Example (2) of Decision Trees Obtained in the Experiment

No.	Rule Element	Table	Layout	Total
0	(root)	89	91	180
1	num_tr ≤ 6	15	82	97
2	border = N	3	69	72
3	img = N	3	18	21
4	img = Y	0	51	51
5	border = Y	12	13	25
6	nest = N	12	5	17
7	nest = Y	0	8	8
8	num_tr > 6	74	9	83
9	nest = N	72	1	73
10	border = N	9	1	10
11	border = Y	63	0	63
12	nest = Y	2	8	10
13	num_tag ≤ 7	0	8	8
14	num_tag > 7	2	0	2

Table 4. Example (3) of Decision Trees Obtained in the Experiment

No.	Rule Element	Table	Layout	Total
0	(root)	86	94	180
1	num_tr ≤ 6	15	84	99
2	num_tag ≤ 12	3	76	79
3	img = N	3	20	23
4	img = Y	0	56	56
5	num_tag > 12	12	8	20
6	nest = N	12	4	16
7	nest = Y	0	4	4
8	num_tr > 6	71	10	81
9	nest = N	69	1	70
10	border = N	10	1	11
11	border = Y	59	0	59
12	nest = Y	2	9	11
13	border = N	0	8	8
14	border = Y	2	1	3

The examples of decision trees shown in Tables 2-4 are typical ones in the 30 trees obtained. The other trees have similar structures from the root node to nodes in depth three as either of the three examples. The 30 trees reveal that the attributes border, num_tr, num_tag and nest is likely to appear in the upper layers of the tree, i.e., these attributes well contribute to the classification.

3.3 Evaluation of Classification Accuracy

Accuracy rates obtained by the three trials of the 10-fold CVs are shown in Table 5. Ten values of the accuracy rates are obtained by a single trial of CV, and the values in Table 5 are the minimum, maximum, mean and SD values of the ten accuracy rates for each trial. In all of the three trials, the maximum rate is 100% so that all checking <table> tags are correctly classified. The mean values are around 90% and the SD values are small, which supports the effectiveness of our method. Improvements for better values of the minimum accuracy rates are the further research challenges.

Table 5. Classification Accuracy Rates (%)

	1st CV	2nd CV	3rd CV
Min.	85	80	80
Max.	100	100	100
Mean	92	89	94
SD	6.0	7.5	6.7

4 Conclusion

In this paper, we proposed a method for detecting layout-purpose <table> tags in webpage HTML sources. The proposed method utilizes a machine learning method for deriving <table> tag classifiers. We have developed a system that utilizes ID3 as the machine learning method. The system derives a decision tree as the classifier from a set of <table> tag data for learning. Classification accuracy was evaluated by 10-fold CVs with 200 webpages collected from the Web. It is found that the purposes can be roughly discriminated with attributes of border, num_tr, num_tag and nest shown in Table 1: these attributes are likely to appear in upper layers in decision trees. In the experiment with the 200 <table> tags collected, mean accuracy rates were around 90%.

In our future work, we'll research whether machine learning methods other than ID3 (e.g., C4.5, multi-layered neural network, support vector machine) can improve the accuracy. These methods will be applied to our system and classification accuracy rates will be compared within the methods.

References

1. HTML 4.01 Specification. W3C Recommendation http://www.w3.org/TR/html4/
2. Web Content Accessibility Guidelines 1.0. W3C Recommendation http://www.w3.org/TR/WCAG10/
3. Ivory, M.Y., Megraw, R.: Evolution of Web Site Design Patterns. ACM Trans. on Information Systems 23(4), 463–497 (2005)
4. Abascal, J., Arrue, M., Fajardo, I., Garay, N., Tomas, J.: Use of Guidelines to Automatically Verify Web Accessibility. Universal Access in the Information Society 3(1), 71–79 (2004)

5. Cooper, M.: Evaluating Accessibility and Usability of Web Sites. In: Proc. of 3rd Int. Conf. on Computer-Aided Design of User Interfaces (CADUI'99), pp. 33-42 (1999)
6. Brinck, T., Hermann, D., Minnebo, B., Hakim, A.: AccessEnable: A Tool for Evaluating Compliance with Accessibility Standards. In: CHI'2002 Workshop on Automatically Evaluating the Usability of Web Sites (2002) http://simplytom.com/research/AccessEnable_workshop_paper.pdf
7. Brajnik, G.: Comparing Accessibility Evaluation Tools: a Method for Tool Effectiveness. Universal Access in the Information Society 3(3-4), 252–263 (2004)
8. Ivory, M.Y., Mankoff, J., Le, A.: Using Automated Tools to Improve Web Site Usage by Users With Diverse Abilities. IT&Society, vol.1(3), pp. 195–236 (2003) http://www.stanford.edu/group/siqss/itandsociety/v01i03/v01i03a11.pdf
9. Ivory, M.Y.: Automated Web Site Evaluation: Researcher's and Practitioner's Perspectives. Kluwer Academic Publishers, Norwell, MA, USA (2003)
10. Scapin, D., Leulier, C., Vanderdonckt, J., Mariage, C., Bastien, C., Farenc, C., Palanque, P., Bastide, R.: A Framework for Organizing Web Usability Guidelines. In: Proc. of the 6th Conf. on Human Factors & the Web (2000) http://www.isys.ucl.ac.be/bchi/publications/2000/Scapin-HFWeb2000.htm
11. Vanderdonckt, J., Beirekdar, A.: Automated Web Evaluation by Guideline Review. Journal of Web. Engineering 4(2), 102–117 (2005)
12. Beirekdar, A., Keita, M., Noirhomme, M., Randolet, F., Vanderdonckt, J., Mariage, C.: Flexible Reporting for Automated Usability and Accessibility Evaluation of Web Sites. In: Costabile, M.F., Paternó, F. (eds.) INTERACT 2005. LNCS, vol. 3585, pp. 281–294. Springer, Heidelberg (2005)
13. Complete List of Web Accessibility Evaluation Tools, W3C WAI http://www.w3.org/WAI/ER/tools/complete
14. TOOLS for Making Interfaces Accessible. in HCI Bibliography http://www.hcibib.org/accessibility/#TOOLS
15. Quinlan, J.R.: Induction of Decision Trees. Machine Learning 1, 81–106 (1986)

The Usability of Accessibility Evaluation Tools

Helen Petrie[1], Neil King[2], Carlos Velasco[3], Henrike Gappa[3], and Gaby Nordbrock[3]

[1] Department of Computer Science, University of York, Heslington,
York YO10 5DD United Kingdom,
[2] DesignedforAll, Prince Consort House, 109 - 11Farringdon Road, London EC1R 3BW
United Kingdom
[3] Fraunhofer-Institut für Angewandte Informationstechnik (FIT)
Schloss Birlinghoven 53754 Sankt Augustin Germany
Helen.Petrie@cs.york.ac.uk, Neil@DesignforAll.com,
carlos.velasco@fit.fraunhofer.de

Abstract. A range of accessibility evaluation tools are available to support Web developers in producing accessible Web resources. It is important that these tools support web developers very well, so that the developers can concentrate on the tasks of Web development. In addition, the tools should enhance Web developers understanding of accessibility issues. This paper presents a usability evaluation of five entry level accessibility evaluations tools. A group heuristic evaluation was conducted, with 5 experts in usability and accessibility working through each tool together, but rating usability problems separately. The results showed that the usability of these tools is surprisingly poor and that they do not support web developers adequately in checking the accessibility of their web resources.

Keywords: Web accessibility, accessibility evaluation, accessibility evaluation tools, usability.

1 Introduction

Testing the accessibility of a website is an important component of its overall development. Testing should include both checking whether the site conforms to the Web Content Accessibility Guidelines (WCAG) [8] and evaluation with disabled people to ensure that they can effectively use the site with different assistive technologies [9].

There are over 100 different tools to assist web developers in testing the conformance of their sites to WCAG [10]. At the most specific level, these include tools which provide information to a developer about a particular accessibility success criterion, but may still require human judgement to decide whether a web resource conforms to WCAG on this success criterion. For example, a tool might display a resource as it would appear to individuals with different colour vision deficiencies; then the developer must check whether the colours used in the resource convey information that is not conveyed by other means (and hence violate WCAG2 Success Criterion (SC) 1.3.2 "Any information that is conveyed by colour is also visually

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 124–132, 2007.
© Springer-Verlag Berlin Heidelberg 2007

evident without colour"). Such a tool is the Accessibility Colour Wheel [4][1]. At the most general level, a tool might conduct a range of automatic tests on a website and inform a developer whether the site conforms to WCAG on success criteria related to these tests. Such a tool is WebXACT from Watchfire [7].

It has been argued [1, 2] that the usability of accessibility evaluation tools should not be considered until the functionality and scope of these tools is clear. Although we agree there are serious questions about the functionality issues, we argue that usability issues always need to be considered from the outset of the development of any interactive system. Studies in human-computer interaction have repeatedly shown that if usability is not considered from the outset, it is extremely difficult to retro-fit and a system that is difficult and frustrating to use is likely to occur [6]. It is important that accessibility evaluation tools are very easy for web developers to use, so that the developers can concentrate on the tasks of Web development and not on how to use the tool and interpret its results.

The accessibility evaluation tools also have an important role in clarifying accessibility issues to developers. Many web developers are still relatively inexperienced about accessibility issues [3], so the more the tools can illustrate how the accessibility evaluation works, the more they will be able to learn in an incidental manner as they interact with the tool. In particular, there are issues that the tools may be conducting a test related to a particular WCAG Checkpoint or SC, but this is not a complete test of the checkpoint or SC. For example, consider Checkpoint 1.1 of WCAG1, "Provide text alternatives for all non-text content ". A particular tool might test this by checking whether each IMG element has an ALT attribute. This is a good initial test, but the fact that a particular IMG element passes this test does *not* mean that it has passed Checkpoint 1.1. It might have a null ALT (ALT = " ") or a meaningless or incorrect description in the ALT attribute (we have found descriptions such as "blah blah blah" and an image of a cat that had a description saying it was an elephant). The tool should make clear exactly what tests it is making, how comprehensive they are and whether the Checkpoint of SC is completely covered by the tests.

To start to investigate these issues, an evaluation of the usability of a range of "entry level" accessibility evaluation tools was undertaken; tools that a novice web developer, who needs a highly usable and informative tool might well use.

2 Method

A group expert evaluation was conducted on five accessibility evaluation tools. The experts worked as a group, spending a maximum of 2 hours using the tool, exploring all the possible functionality. Any expert was allowed to propose a usability problem about the tool, but then all the experts privately noted whether a proposed problem was indeed a problem, and if they thought it was a problem, rated its severity. Thus different experts might identify different numbers of problems for a particular tool, and rate them differently.

[1] Examples of products, commercial or otherwise, are provided only as illustrative of particular points, and are not endorsements of these products.

2.1 Participants

Five experts in both the usability and accessibility fields participated. 2 were female, 3 were male. Ages ranged from 31 to 49. Years of experience in the usability field ranged from 2 to 15 years. Years of experience in the accessibility field ranged from 1 to 13 years.

2.2 Tools Evaluated

The evaluations were conducted in July 2005. Tools may have changed since then. The following information describes the tools when they were tested.

Five evaluation tools were evaluated:

- Site Valet (http://valet.webthing.com/)
- WAVE 3.0 (http://www.wave.webaim.org/)
- Cynthia Says (hhtp://www.cynthiasays.com)
- Bobby 5.2 (www.watchfire.com/default.aspx)
- Deque Ramp (http://www.deque.com/products/ramp/index.php)

2.2.1 Site Valet by WebThing

The online service of Site Valet offered by WebThing evaluates a single web resource. The service claims to check the page according to WCAG2 (this seems

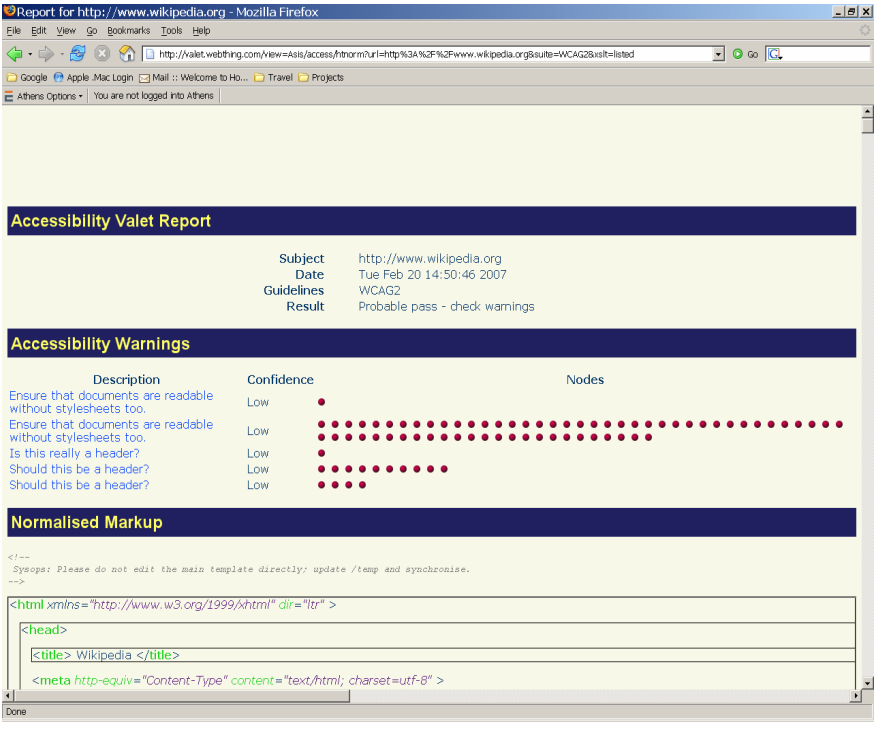


Fig. 1. Site Valet accessibility reporting page

extremely unlikely, although it has no bearing on the results of the present evaluation). It is not possible to customize the coverage of the accessibility check, so a page cannot be inspected for conformance to other accessibility standards such as WCAG 1.0, Section 508 or particular priority levels of these guidelines. No information about the scope of the service, terms of use or other such information is provided. The service is free.

The Site Valet accessibility reporting page is divided into three parts (see Figure 1). The first provides a summary of what standards have been checked, the URL of the Web page evaluated and a result in terms of "pass" or "fail". The second part displays all relevant accessibility warnings with a description, for example "table cells should be associated with headers", a column indicating the level of confidence about this warning (high, medium, or low) and the instances of this accessibility warning indicated by a red circle under the header "node". The nodes are directly linked to the occurrence in the markup which is displayed in the third part of the report.

2.2.2 Wave 3.0 by WebAIM

The WAVE Accessibility Tool Online Web Service was originally developed at Temple University Institute on Disabilities in Pennsylvania and has now been taken over by WebAIM (Web Accessibility in Mind) at the Center for Persons with Disabilities (CPD, Utah State University).

The tool is free of charge and is intended for personal, non-commercial use. The terms of use state that the tool is offered as an "as is" product, meaning no warranty is provided and that compliance to a certain Web accessibility standard cannot be ensured by usage of the service. A navigation link leading to the terms of use is displayed clearly next to the submit button for an accessibility evaluation on the home page.

WAVE evaluates a single Web resource. The accessibility reporting page displays the tested Web resource with icons indicating problems on the resource in the following colours (see Figure 2):

- red icons for errors that need repair
- yellow icons for alerts that need to be checked for possible errors
- green icons for accessibility features that need to be checked for accuracy
- blue icons for structural and semantic elements that may support accessibility and should be checked for accuracy

The meaning of the icons is explained in a separate document which can be opened from the accessibility reporting page in a new window. In this document, the relevance of each icon in relation to WCAG 1.0 and Section 508 is explained.

The reporting page does not provide an evaluation result in terms of "pass" or "fail". WAVE allows the user to customize the accessibility check for WCAG1.0 Priorities 1, 2 and 3, and Section 508, the elements to be displayed in the report (e.g. tables, images, frames) and special WAVE features such as the implications of linearised reading order by an arrow to the right, borders on all tables, etc.

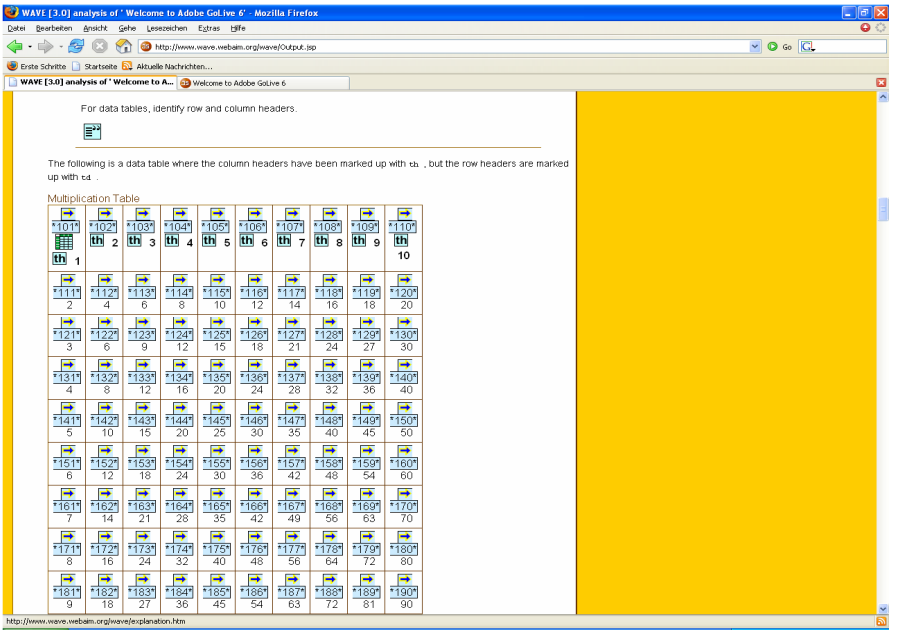


Fig. 2. Wave 3.0 accessibility reporting page

2.2.3 Cynthia Says by HiSoftware

Cynthia Says is a single web resource evaluation tool from HiSoftware. As the company states, Cynthia Says is designed to offer a service for all users who wish to create accessible websites, including those with little or no knowledge. The section "about this site" explains that Cynthia Says evaluates a web resource against tests for Section 508 and WCAG1 by employing HiSoftware's AccMonitor Server technology and conducting accessibility verification on its central server with the result being sent back to the user's browser. This service is offered free of charge.

HiSoftware states in its terms of use that Cynthia Says shall only be used for personal, non-commercial purposes. It is offered on an "as is" basis, meaning that no guarantee for accuracy of the results is given. It is also stated that the test results need to be reviewed by a human and that manual checks as listed in the accessibility report need to be performed.

On the accessibility reporting page, the resource that has been evaluated is indicated and whether it has passed or failed automated verification. Result are presented in the form of the tabular summary of WCAG1 checkpoints. Below each checkpoint the tested that have been evaluated are listed with the result returned by Cynthia Says: "Rule: 1.1.2 - All INPUT elements are required to contain the alt attribute or use a LABEL" with the result "No invalid INPUT elements found in document". In the column next to this statement, fulfilment of this checkpoint is rated. This can be a "Yes", "No", "Not Verified" (N/V), "No related elements were found in the document" (N/A) or a "Warning". Cynthia Says also provides the line number in the source code of the tested page where a problem has been detected.

2.2.4 Bobby 5.2 by Watchfire

The Bobby 5.2 is the desktop version of the well-known Bobby tool and is designed to conduct accessibility audits encompassing an entire website. It is a commercial product that aims to identify barriers to accessibility in accordance with WCAG1 and Section 508 guidelines. It spiders through a website page by page conducting 90 accessibility checks per page. Several types of accessibility reports are offered. There is a summary statement that gives an overview of all pages reviewed with the status (pass/fail) and the number of accessibility errors and possible accessibility issues. It offers the possibility to link to a particular page to access an accessibility report for that page. The individual page report follows the same outline as the summary report. The only difference is that here for some of the listed accessibility errors and issues, a line number is provided indicating the location in the source code where they have been identified. All accessibility findings specified in the reports are linked to a brief explanation on what is meant by this finding, how this issue can be avoided, and a rationale for this finding including which user groups are affected. At the end of the explanation a link to the according reference in WCAG1 and Section 508 is offered.

2.2.5 Deque Ramp

Deque Ramp is a commercial web accessibility tool that brands itself as being both 'powerful and comprehensive' to sighted and non-sighted web developers. Deque Ramp promotes itself as a system that is platform independent and designed to test and retest a site's accessibility.

Deque Ramp evaluates a website either page-by-page, or by spidering through the website as dictated by the author of the evaluation. This enables developers and analysts to enter the website's URL, and then select how deep into the website they wish to test. Unlike basic one-page tools, Deque Ramp states it can navigate the most complex web sites and produce highly usable reports. The tool produces an array of reports including an overview of problems that directly relate to missing images or code violations in one click and more through in-depth reports relating to checkpoint violations.

The tool also includes a 'Wizard' function, to assist the web developers in displaying the results relevant to their particular criteria.

2.3 Procedure

Each expert was provided with a pile of Usability Problem Report Sheets and a numbered list of Nielsen's heuristics (Nielsen and Mack, 1994). Prior to each evaluation session, one member of the group of experts familiarised themselves with the tool to be evaluated. This person led the group through a number of accessibility audits using the tool and a general exploration of the tool's features. The group worked its way through the tool systematically as far as possible, looking for all potential usability problems.

Each time a potential usability problem was discovered (and all the experts were encouraged to propose such problems), a description of the potential problem was agreed between the experts and noted on the Usability Problem Report Sheets. Each expert then privately rated whether he/she thought it was a problem or not. Thus it was not necessarily the case that all the experts would have the same number of

reported usability problems, as they had the option of indicating that they did not think a proposed problem was actually a problem. If they did indicate that it was a problem, they also noted the severity of the problem using the 4 point scale proposed by Nielsen and Mack (1994): usability catastrophe, major usability problem, minor usability problem, cosmetic usability problem. They also indicated which of Nielsen's usability heuristics would account for the problem, if any.

This procedure was repeated until the experts were confident that all the usability problems in the tool had been uncovered. Each evaluation took approximately 2 hours to complete.

3 Results

3.1 Results for Site Valet

For Site Valet, a total of 16 usability problems were identified. These consisted of 1 catastrophe, 12 major, 2 minor and 1 cosmetic usability problem. Table 1 provides a summary of the catastrophe problem found in Site Valet (space permits us only to list catastrophic problems for each tool).

3.2 Results for Wave

For the Wave 3.0, a total of 12 usability problems were identified. These consisted of 7 catastrophe, 4 major, and 1 minor usability problems. Table 2 provides a summary of the most critical usability problems found on the Wave. This tool generated the least number of unique usability problems of all those tested, but over half of those issues identified were awarded the highest mean severity rating (catastrophe).

Table 1. Catastrophic usability problem found with Site Valet

Description	Mean severity rating	Relevant heuristics
Functionality of buttons "Accessibility", "Validate" and "links" are unclear – does "Accessibility" refer to WCAG P1, P2 and P3 or Section 508?	3.6 Catastrophe	Clarity of information

3.3 Results for Cynthia Says

For Cynthia Says, a total of 16 problems were identified, consisting of 13 major and 3 minor usability problems. This tool generated no catastrophe usability issues, but the majority of problems (81%) were defined as major usability problems. The experts also found it particularly hard to agree on which heuristics were relevant to these problems.

3.4 Results for Bobby 5.2

For Bobby 5.2, a total of 19 usability problems were identified. These consisted of 1 catastrophe, 10 major, 6 minor and 2 cosmetic usability problems. Table 3 provides a summary of the catastrophe problem found in Bobby 5.2.

Table 2. Catastrophic usability problems found with Wave 3.0

Description	Mean severity rating	Relevant heuristics
Tool does not provide a fluid route to effectively compare results and source code	3.8 Catastrophe	Lack of task support
No summary overview of test results e.g how many errors, what level the page has reached (A, AA etc)	3.8 Catastrophe	Lack of task support
Icons are non-intuitive	3.8 Catastrophe	Match between system/real world
Information explaining the icons is unclear	3.8 Catastrophe	Help and documentation
Lack of clear mapping between results page and actual web page	3.6 Catastrophe	Match between system/real world
Terminology used for preferences is unclear	3.6 Catastrophe	Match between system/real world
Nothing about what guidelines/level checked against	3.6 Catastrophe	Help and documentation

3.5 Results for Deque Ramp

For Deque Ramp, a total of 21 usability problems were identified. These consisted of 1 catastrophe, 16 major, and 4 minor usability problems. Table 4 provides a summary of the catastrophe problem found in Deque Ramp.

Table 3. Catastrophic usability problem found with Bobby 5.2

Description	Mean severity rating	Relevant heuristics
Disassociation between hitting “scan now” button and system status at bottom of page	3.6 Catastrophe	Visibility of system status

Table 4. Catastrophic problem found with Deque Ramp

Description	Mean severity rating	Relevant heuristics
No relationship between “interactive checks” and “violations”	3.8 Catastrophe	Consistency and standards

4 Discussion and Conclusion

The results of this evaluation are disappointing. All the tools evaluated had considerable numbers of very obvious and serious usability problems that could easily have been avoided. The home page for these tools requires only a field to enter a url, information about the tests to be conducted and possible options to choose, but even there problems often occurred. It was often not clear whether only the single resource was being tested, or child resources as well. The experts were sometimes forced to

conduct test evaluations to establish this. In addition, it was not always clear what guidelines/levels of guidelines would be tested.

But the bulk of problems occurred in the accessibility reporting pages. As one expert commented "if I was unsure about accessibility before I started using this tool, I would be completely baffled after using it". The presentation of results was almost always unclear. Information such as what tests had been made, what guidelines/levels of guidelines these related to, what priority level a resource had passed, should all be very clear to a user when viewing an accessibility report, but was rarely the case. Experts felt that web developers would want to move between information about the overall conformance levels, specific accessibility violations, the code where the violations occurred and information about how to repair that code. Several of the tools have begun to made attempts to provide those links, but there were many usability problems in this area.

Overall, these entry level accessibility evaluation tools are not making the vital task of checking the accessibility of web resources easy for web developers. Those who are experts in web accessibility need to work to provide better support for the wider community of web developers to encourage them to consider accessibility in their development process.

References

1. Brajnik, G.: Automatic web usability evaluation: what needs to be done? In: Proceedings of 6th Conference on Human Factors and the Web. Austin, Texas. Available at (2003) http://www.tri.sbc.com/hfweb/brajnik/hfweb-brajnik.html
2. Brajnik, G.: Comparing accessibility evaluation tools: a method for tool effectiveness. Universal Access in the Information Society 3(3 – 4), 252–263 (2004)
3. Disability Rights Commission, The Web: Access and Inclusion for Disabled people. London: The Stationery Office (2004)
4. Mazzocato, G.: Accessibility Colour Wheel (Accessed: 15/02/07) Available at: http://gmazzocato.altervista.org/colorwheel/wheel.php
5. Nielsen, J., Mack, R.L.: Usability inspection methods. John Wiley, New York (1994)
6. Preece, J., Rogers, Y., Sharp, H.: Interaction Design: Beyond Human-Computer Interaction. John Wiley & Sons, New York (2002)
7. Watchfire. WebXACT (Accessed: 15/02/07) Available at: http://webxact.watchfire.com/
8. Web Accessibility Initiative. Web Content Accessibility Guidelines Overview (Accessed: 15/02/07) Available at: http://www.w3.org/WAI/intro/wcag.php
9. Web Accessibility Initiative. Involving Users in Web Accessibility Evaluation (Accessed: 15/02/07) Available at http://www.w3.org/WAI/eval/users.html
10. Web Accessibility Initiative. Complete list of web accessibility evaluation tools (Accessed: 15/02/07) Available at: http://www.w3.org/WAI/ER/tools/complete

Accessibility Works: Enhancing Web Accessibility in Firefox

John T. Richards, Vicki L. Hanson, Jonathan P. Brezin, Calvin B. Swart, Susan Crayne, and Mark R. Laff

IBM T. J. Watson Research Center, 19 Skyline Dr, Hawthorne, NY, USA
{ajtr,vlh,brezin,cals,crayne,mrl}@us.ibm.com

Abstract. This paper reviews several techniques we have discovered while trying to extend the Firefox browser to support people with visual, motor, reading, and cognitive disabilities. Our goal throughout has been to find ways to make on-the-fly transformations of Web content including adjustments of text and image size, text style, line and letter spacing, text foreground color, text background color, page background removal, content linearization, and reading text aloud. In this paper, we focus primarily on the changes we make to the browser's Document Object Model (DOM) to transform Web content. We review the kinds of approaches we have used to make DOM modifications sufficiently fast and error free. We highlight the problems posed by Web pages with a mix of static and dynamic content generated by client-side scripts and by Web pages that use both fixed and relative placement of page elements, pages of the sort we expect to see in increasingly in the future.

Keywords: Web accessibility, transcoding, Document Object Model.

1 Introduction

The W3C has developed guidelines for accessible Web pages [1] [2]. Websites that conform to these guidelines provide a number of critical features allowing blind or deaf users to get Web information in another modality. These guidelines, however, often fall short of enumerating the kinds of modifications needed by users with what might be viewed as less severe deficits (e.g., enabling low-vision users to use the vision they do possess to greater advantage, or simplifying page structure to help users unable to navigate complex layouts). Even if the guidelines were broader, they would not solve the problem, because many websites do not conform to them, and the cost of coming into conformance is prohibitive. Even conforming websites do not typically allow individual users to control their browsing experience as outlined in the User Agent work of the W3C [3] [4]. Allowing individuals to adapt Web content for their unique needs has inspired our work on Web page transformations.

For the past several years, we have been developing extensions to the Internet Explorer™ (IE) browser to support on-the-fly transformation of Web content to support Web access for people with limited vision and dexterity, as well as for people with language or cognitive issues that impact their ability to use the Web. The key

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 133–141, 2007.
© Springer-Verlag Berlin Heidelberg 2007

idea was to provide a single software application that could serve a wide number of users with a variety of needs.

Our early work explored using a proxy server to transform a Web page's source before it got to the browser, but a number of problems with this approach motivated us to move instead to all client-side transformations [5]. The resulting transformations were based on a number of techniques including: modification of the Windows^{tm} system registry, dynamic generation of user style sheets, manipulation of the browser's internal model of the loaded Web page, and the launching of separate executables for filtering user input and magnifying images. This Internet Explorer implementation, previously reported [6] [7], is available in seven languages and used by thousands of people worldwide.

Two years ago, we began a fundamental redesign of this software for use in the Firefox browser [8]. Part of the appeal of Firefox is that it runs on multiple platforms including Windows, Mac OS, and Linux. While we use Mozilla's XPCOM mechanism [9] to access platform specific speech functions, and while our mouse and keyboard adaptations require platform specific interaction with the operating system [10] [11], we concentrated most of our Firefox work on changing the browser's internal model of the loaded Web page – its Document Object Model (DOM). In this paper, we discuss our approach and some of the techniques we used to achieve enhanced Web accessibility for users of Firefox.

2 Changing Page Elements

We have created several Firefox accessibility transformations for changing the appearance of Web pages. One useful transformation allows foreground and background colors to be changed to heighten contrast, or produce a custom color combination for people with certain types of low vision or dyslexia. Another lessens static clutter by removing background images. A third eliminates dynamic distractions by stopping the movement of animated gifs. What these transformations have in common is that they do not change the placement of elements on the page since visual elements retain their sizes and locations.

A more complex class of accessibility transformations involves changing the size of page elements. Transformations found to be useful from our earlier IE implementation include magnifying the entire page, changing the spacing between lines of text, changing the spacing between letters of text, and changing the font size of text.

We have found that magnifying the entire page, easily done in IE by dynamically generating a user style sheet with the "zoom" attribute, is impractical in Firefox. Since at least earlier versions of Firefox lack a native zoom feature we would be forced to re-layout the entire page, element by element.

Unlike zooming, the line spacing and letter spacing, transformations can be accomplished straightforwardly in Firefox. The line space transformation is made by modifying the page body's line height style along with attributes of certain TABLE, DIV, PARAGRAPH, and LIST items in the DOM. The complete set of elements requiring attribute modification has been determined somewhat empirically so may not cover all current and future Web page designs. But it seems to work well for most.

Letter spacing is somewhat simpler than line spacing in that we can simply modify the page body style for letter spacing. We have not yet encountered pages that modify this style at the element level so we can bypass the element modifications needed for line spacing.

Changing text size is trickier than might be supposed. There is support in HTML itself for asserting *relative* text size, and for those situations, changing the default size scales the text nicely. But the range of enlargement is too small for many of our users since only four larger sizes are available. To go beyond this range we have to use explicit font size specifications, and to enforce them, we have to process all style sheet rules and FONT elements in the DOM. But font size changes can easily interact with explicit placement of nearby elements, causing text that was adjacent to another element to overlap. If pages include fixed width elements the problem can quickly get out of hand as shown in Figure 1.

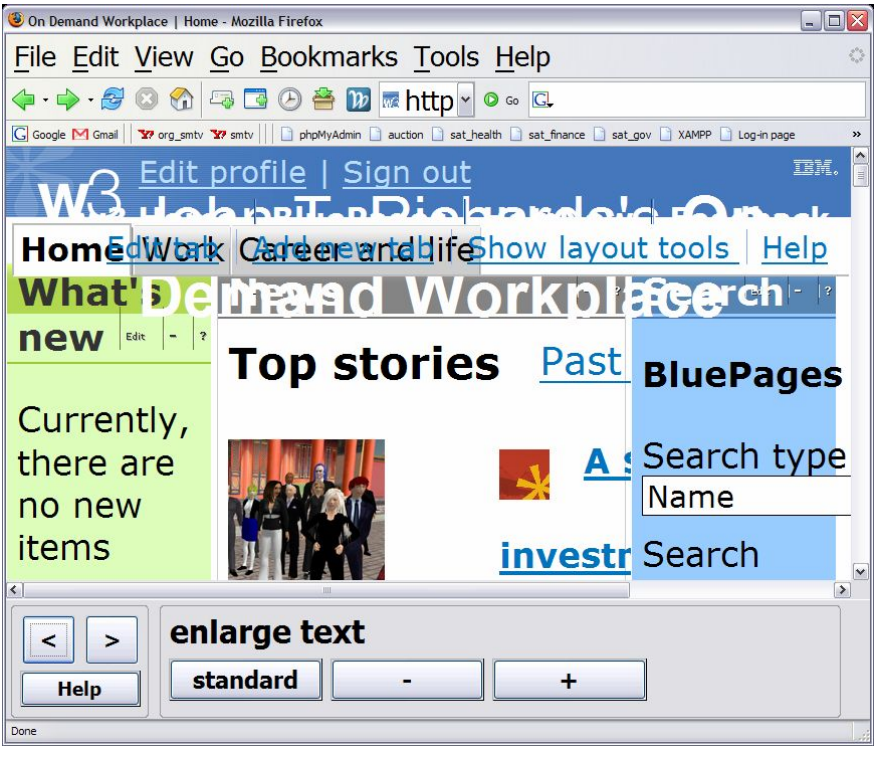


Fig. 1. Potential problems with naïve text enlargement

3 Changing Page Geometry

Our software is intended to serve people with a variety of disabilities. In an individual person these disabilities often interact. For example, a person may need a page to be

enlarged so as to be easily read, but may also need to minimize use of the mouse due to a motor disability.

Consider the case of a multicolumn page. As it is enlarged, it can extend beyond the right hand edge of the browser window, causing a horizontal scroll bar to appear. Horizontal scrolling, which is already undesirable for those without any motor disability, can become a major barrier to effective browsing for those who find mouse and keyboard operation difficult. To accommodate this need, we created a feature that causes multiple column content to be 'linearized' into a single column as shown in Figure 2.

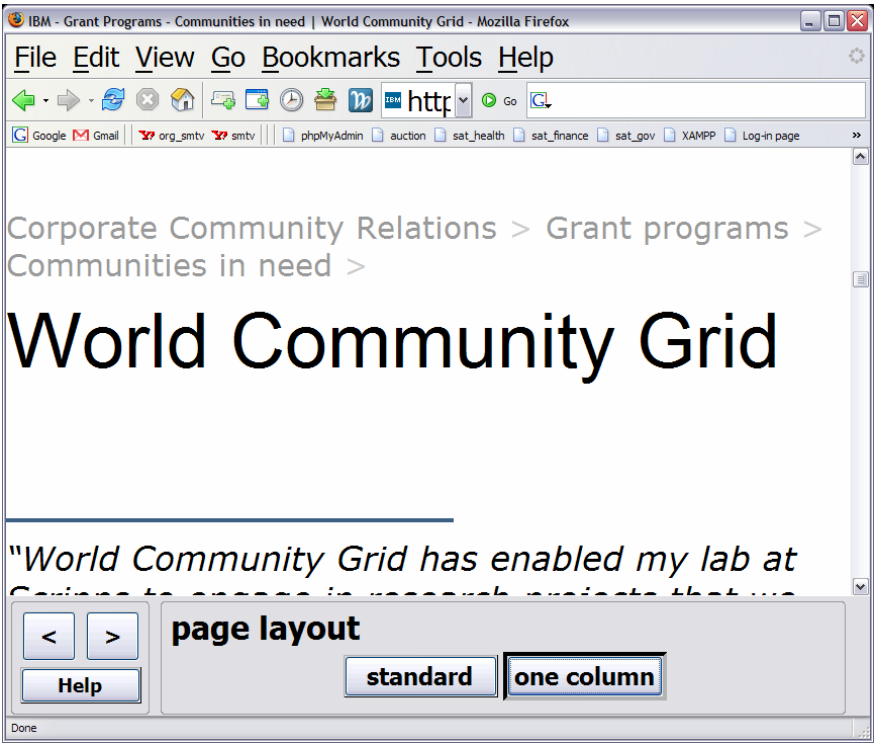


Fig. 2. A page after substantial text enlargement and linearization

Like the previous size manipulations, page linearization is changing the geometry of the page. But the change here is so extensive that we cannot affect it by merely manipulating attributes of individual elements. We must actually modify the overall structure of the DOM. And we must do this in a way that maximizes the likelihood that scripts and style rules attached to the document will continue to operate correctly. For example, making each TD or TH element its own row by inserting TR elements is a very bad idea: just imagine that the original TR had an id like "PopulationTableHeadings" that a style rule depended on. Either we recognize the dependence and replace it with a dependence on a class (which we have to create), or we have to maintain the containment hierarchy implied by the original DOM. A better

solution is to replace every table element with a DIV that has all the relevant attributes of the element copied into it. That way the parent-child relations in the DOM are unchanged. Of course, we still have to take into account the border implied by the TABLE element (and possibly with the table as well).

4 Banner Text and Spoken Text

For some users, the most effective transformation of Web content is to greatly enlarge small blocks of text and/or to read portions of text aloud using text-to-speech software. Reading text aloud has proven to be especially useful for people with very low vision and those unable to read for reasons of either language competency or general reading ability [6] due to dyslexia, low literacy, or cognitive disability. Using IBM's ViaVoice[tm] product, our software will read aloud selected portions of a page or unselected text regions up to the whole page. Assuming that multiple ViaVoice languages are installed on a user's computer, the software can even read a page with the correct language pronunciation. It does this by using various techniques for determining the language of the page and then using the appropriate ViaVoice language. In the simplest case the language is specified by the <HTML lang="…"> attribute [12]. Relatively few pages, however, follow this W3C recommendation, so other techniques, based on an analysis of language specific characteristics of the page content, must be used.

Blind users, of course, rely exclusively on spoken text provided by screen readers. Such screen readers require the user to memorize a large number of specialized keyboard commands to control navigation and speech output. While the payoff for dedicated users is obvious, casual users will not memorize complex keyboard commands. Moreover, some cognitively disabled users would not be able to memorize these commands. Fortunately, our target users can all be assumed to have at least some usable vision. The technique we have evolved – something you could call "hover to hear" – takes advantage of this fact. By bring the mouse pointer to rest above the text area of interest our users can see greatly enlarged text and/or hear text read aloud from the beginning of the corresponding text element.

To support this "hover to hear" technique we must first determine what the user is pointing at. In the simple case, the mouse is over a relatively small text node and we can start reading from the beginning of that node. For larger blocks of text it is difficult to determine where in the text node the mouse pointer lies. As such, we would be forced to read from the beginning of the block even though the user is pointing somewhere within it. As you might imagine, this is quite frustrating for our users. To make spoken text more usable, we insert SPANs into the DOM to break up large text blocks into smaller units allowing finer localization. While in principle this should affect neither scripts nor style rules, in practice it can result in the text and explicitly positioned elements being drawn incorrectly vis-à-vis one another. The general problem of how to deal with text within explicitly positioned elements remains open.

Assuming we can find the starting point in the text corresponding to the user's expectations (and we usually can), we try to keep reading until the user asks us to stop by moving the mouse outside the text area, or clicking, or hitting any key. It has also

proven desirable to dynamically highlight what is being read, even over page boundaries (forcing us to scroll if necessary). This is fairly straightforward as long as one can continue to scroll down. But, if the next text in document order is above the current text in the display, we have found it is best to stop, because scrolling the display back may leave the user bewildered as to where the current text is on the screen.

Of course, in order to do any of this work, our software must see suitably localized mouse as well as keyboard events. In the case of IE, we ran a thread parallel to the browser thread that attached event handlers to every DOM node that we cared about (which, alas, were most nodes). In this way, we could get control and react to the mouse moving over a piece of text and hovering there. We were forced to do this at the element level in IE because one cannot tell from the Event object where in the DOM the event occurred, only that it occurred inside whatever element handled the event. Firefox's Event API provides a bit more information about where the event occurred relative to the DOM so we only need event handlers for the root (body) element and each FRAME or IFRAME. The reason for the latter is that when we are trying to read the "next" text, we sometimes have to go to the document containing the document we are now in (the parent frame of our current frame). Hence we need to track what frames have the text we are reading.

Handling the mouse events involves a little finesse. By trapping mouse-up events one can track selection, and (assuming the user has told us to) we can read the selection. Click events, on the other hand, always end the current reading session and may be part of a double click that selects a word. A simple but not trivial state machine handles this bookkeeping. We also use a timer to know when to start reading the next text element. Once speaking starts, we start a timer, and when it fires, we ask whether we are still speaking (unfortunately there is no end-of-spoken-text event available to us). If we are done, we proceed to find the next block of text to be read and do so. If not, we restart the timer.

5 Dynamically Constructed DOMs

A browser's DOM is a fairly straightforward data structure. As specified by the W3C [13], the DOM represents hierarchical HTML documents basically as trees and provides element access and manipulation through an API intended to be conveniently used by script writers. As a practical matter, all DOM renderers allow an HTML DOM to be "live". That is, it is automatically re-rendered whenever any change is made to it that might affect the rendered appearance. One common way to make such a change is to create a piece of ordinary HTML as text, and then to use the scripting API to insert this text as the content of an element already present in the DOM. Since scripts using this technique can run quite some time after a document is loaded it is by no means clear when a DOM is *finally* ready to be modified for purposes of enhancing the readability, or general appearance, or overall layout of the document.

Consider the following simple example. Figure 3 shows a help page that is provided with our Firefox software. In it we describe how text foreground and background colors can be changed to heighten overall contrast or provide custom color combinations suitable for people with some types of dyslexia. This help page

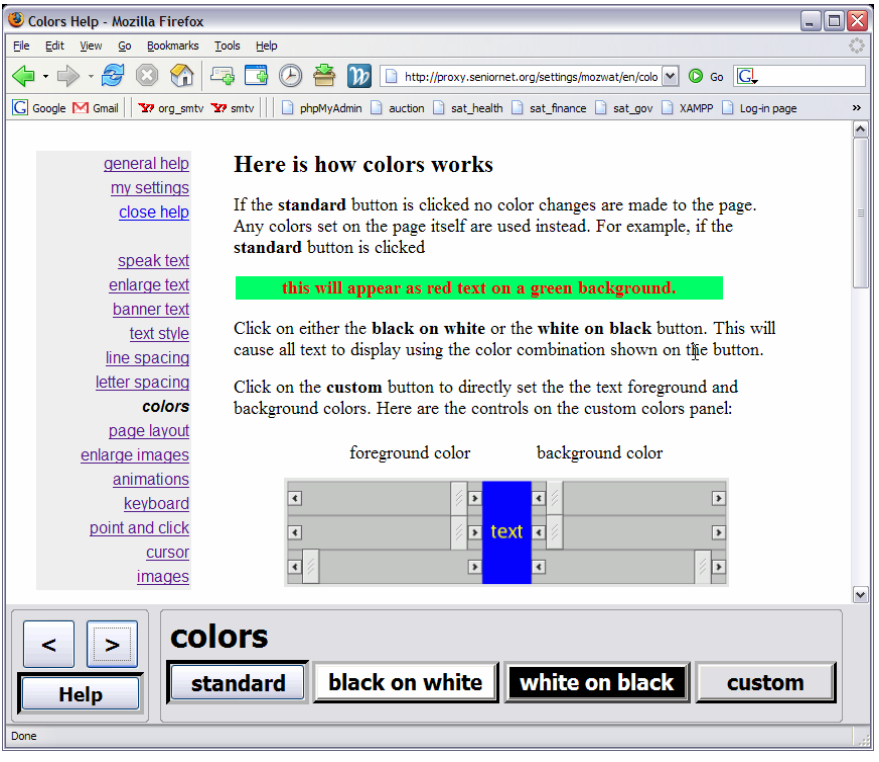


Fig. 3. Help page for changing colors with dynamically constructed navigation bar on the left

also has a navigation bar on the left hand side to allow users to conveniently move to help for the other transformations we provide.

Now consider how this fairly normal looking navigation bar is actually created. In the early days of our work, it was simply coded as part of the static HTML of each individual help page. This worked fine as long as the number of pages was small and their order didn't change too frequently (the order corresponding to the order of transformation panels as one clicks the left/right arrow buttons on the bottom panel). But as the number of transformation panels grew, as we started responding to feedback from our users as to their preferred ordering of panels, and as we started providing translations for multiple national languages, the time required to make even a simple help change became prohibitive. To remedy this, the help pages were modified so the navigation bar became a placeholder, an empty DIV, whose "inner HTML" was set by a script in the onLoad handler of each help page. This script iterates over an ordered array of national-language independent help page identifiers, gets the translated string for the current user's language's name for that page, and adds it as a table row either as a clickable anchor or as an emphasized, italicized string for the current help page.

While this script runs without perceptible delay, the insertion of the inner HTML into the DIV (and hence into the underlying DOM) came too late for our extension to see it in its DOM traversal, this traversal being similarly triggered by the onLoad

event. As such, the page would display with the navigation bar having the original gray background rather than the background dictated by any color transformation the user may have set. Since the entire help system was designed to dynamically illustrate how each transformation would be applied to pages on the Web, this was an unacceptable outcome.

To fix this anomaly, we did two things. First, we optimized the script building the navigation bar's inner HTML so it ran in the shortest possible time. Next we delayed our DOM walk by a fraction of second to allow the full DOM to be built. It does not take much imagination to see that this is not a very robust solution. But lacking a dependable event that tells us when a page's onLoad scripts have finished executing, it may be the best we can do.

In general, we are finding that more and more Web content is being constructed using dynamic techniques of the sort illustrated by this simple example. The increasing prevalence of the software pattern known as Asynchronous JavaScript and XML (AJAX) makes the situation even more challenging since a DOM can change at any time in response to user events such as clicking on a button or selecting an item from a pull down list. If we are to continue to make Web content accessible we need to develop a range of software techniques that can cope with this sort of dynamicity [14] [15].

6 Conclusions

Enhancing Web accessibility through on-the-fly content transformation remains challenging. In part this is due to the increasing richness of techniques employed by content designers. In addition, the use of AJAX and client-side, script-based content generation, while having the desirable effect of heightened interactivity and ease of use, make DOM manipulation a complex and often multi-step process. The use of style sheet rules to control element placement further complicates even simple transformations such as font size adjustments.

Between the occasional browser bug, the arguably misguided use of fixed or relative element placement, and the sheer complexity of some commercial web sites, we are faced with enough unpredictability that our work becomes essentially empirical. We can be guided by how things should work, but are often forced to create workarounds for how things actually work. Of course, empirically testing against the enormity of the Web is not feasible. But even with these difficulties our transformations have attained a level of speed and accuracy that allows them to be used on a daily basis by many disabled people around the world.

That being said, there are a few things that would make our task easier. IE, for example, would benefit from the sort of continuous font scaling that Firefox supports. And Firefox would be easier to augment if we could "batch" changes to a style sheet so that multiple changes would cause only single re-rendering. Both browsers could also provide new APIs to allow us to obtain the sorts of fine grained location-to-structure mappings we now have to construct in a laborious and sometimes error prone way. For example, all browsers know what element lies beneath the current pointer. It would be quite useful to have an API that could directly return the text offset corresponding to the pointer location to better support tasks like reading text aloud. More generally, APIs that caused the browser's renderer to make the sorts of

transformations we have found useful for people with visual disabilities would be highly desirable.

Finally, a richer set of events could be signaled (or at least the current set of events could be more reliably signaled) to allow browser extensions to know when it is safe to begin manipulating the DOM. While our software works reasonably well with the current set of events and the current mix of Web content, the increased interactivity associated with "Web 2.0" will likely require additions to the event model if accessibility transformations such as ours are to continue to work well.

References

1. Web Accessibility Initiative (Retrieved November 19, 2006) from http://www.w3.org/WAI/
2. Web Content Accessibility Guidelines Working Group (WCAG WG), (Retrieved February 15, 2006) from http://www.w3.org/WAI/GL/
3. Gunderson, J.: W3C user agent accessibility guidelines 1.0 for graphical Web browsers. Universal Access in the Information Society 3, 38–47 (2004)
4. Jacobs, I., Gunderson, J., Hansen, E. (eds.): W3C user agent accessibility guidelines 1.0. (Retrieved November 19, 2006) (2002) from http://www.w3.org/TR/UAAG10
5. Hanson, V.L., Richards, J.T.: A Web Accessibility Service: An Update and Findings. In: Proceedings of the Sixth International ACM Conference on Assistive Technologies, ASSETS 2004, pp. 169–176. ACM, New York (2004)
6. Hanson, V.L., Richards, J.T.: Achieving a usable World Wide Web. Behaviour and Information Technology 24(3), 231–246 (2005)
7. Hanson, V.L., Snow-Weaver, A., Trewin, S.: Software personalization to meet the needs of older adults. Gerontechnology 5(3), 160–169 (2006)
8. Hanson, V.L., Brezin, J., Crayne, S., Keates, S., Kjeldsen, R., Richards, J.T., Swart, C., Trewin, S.: Improving Web accessibility through an enhanced open-source browser. IBM Systems Journal 44(3), 573–588 (2005)
9. XPCOM. (Retrieved November 19, 2006) from http://www.mozilla.org/projects/xpcom/
10. Trewin, S.: Automating accessibility: the dynamic keyboard. In: Proceedings of the 6th international ACM SIGACCESS Conference on Computers and Accessibility (Atlanta, GA, USA, October 18 - 20, 2004), pp. 71–78. ACM Press, New York (2004)
11. Trewin, S., Keates, S., Moffatt, K.: Developing steady clicks: a method of cursor assistance for people with motor impairments. In: Proceedings of the 8th international ACM SIGACCESS Conference on Computers and Accessibility (Portland, Oregon, USA, October 23 - 25, 2006), pp. 26–33. ACM Press, New York (2006)
12. W3C HTML 4.01 Specification (Retrieved November 19, 2006) from http://www.w3.org/TR/html4/struct/dirlang.html
13. W3C Document Object Model (DOM) Level 1 Specification (2nd edn.) (Retrieved November 19, 2006) from http://www.w3.org/TR/2000/WD-DOM-Level-1 - 20000929/ Overview.html
14. Gibson, G.: AJAX accessibility overview (Retrieved November 19, 2006)(2006) from http://www-306.ibm.com/able/resources/ajaxaccessibility.html
15. Gibson, B., Schwerdtfeger, R.: DHTML accessibility: solving the JavaScript accessibility problem. In: Proceedings of the 7th international ACM SIGACCESS Conference on Computers and Accessibility (Baltimore, MD, USA, October 09 - 12, 2005). Assets '05, pp. 202–203. ACM Press, New York (2005)

Improving Accessibility for Existing Websites Spanning Multiple Domains

Takashi Sakairi[1], Takuya Ohko[2], and Hidemasa Muta[1]

[1] IBM Japan, Tokyo Research Laboratory
[2] IBM Japan, Accessibility Center
1623-14, Shimo-tsuruma, Yamato, Kanagawa 242-8502, Japan
{sakairi,ohkot,hmuta}@jp.ibm.com

Abstract. There are demands for improving accessibility in existing websites by enlarging text and changing the colors. Typical solutions use technologies such as ActiveX that can only run on a specific client environment. JavaScript is supported in many Web browsers, and it can be used to add new functions for improving the accessibility of existing websites. However, Web browsers prohibit JavaScript from accessing webpages of another domain, so it is difficult to improve accessibility for related websites spanning multiple domains. This paper describes a method that solves the problem.

Keywords: Accessibility, Websites, Dynamic HTML, JavaScript, Same Origin Policy.

1 Introduction

The Web plays important roles in daily life and business. Many organizations develop websites for various purposes, such as to deliver information and for electronic commerce. Meanwhile, the ratio of senior citizens is increasing and will soon be increasing quite rapidly in many countries. However, there are many obstructions to senior citizens' access to the Web. Due to the complications of weak eyesight, many seniors are uncomfortable about browsing webpages crammed full of information and using small, hard-to-read fonts. In response, many websites' owners recognize that accessibility is increasingly important [1, 2, 3]. It is possible to improve accessibility for people with limited vision by enlarging the text and by changing the colors.

Fig. 1 shows two types of accessibility tools. The first type is an accessibility tool for a Web browser (a). Users install such an accessibility tool by themselves [4]. Users can access any Web server with the tool, but it is difficult for novice users to install such tools. The second type is an accessibility tool for a Web server (b). Providing accessible websites without troubling the user is important for many organizations, especially for public institutions. This paper describes an accessibility tool for a Web server.

When a website's owner decides to improve accessibility, it is not always the case that the affected webpages are in a single domain. For example, a local government may have a website, http://www.city.jp/, but a university established by the local government may have a separate website, http://www.u-city.ac.jp/, in another domain.

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 142–149, 2007.
© Springer-Verlag Berlin Heidelberg 2007

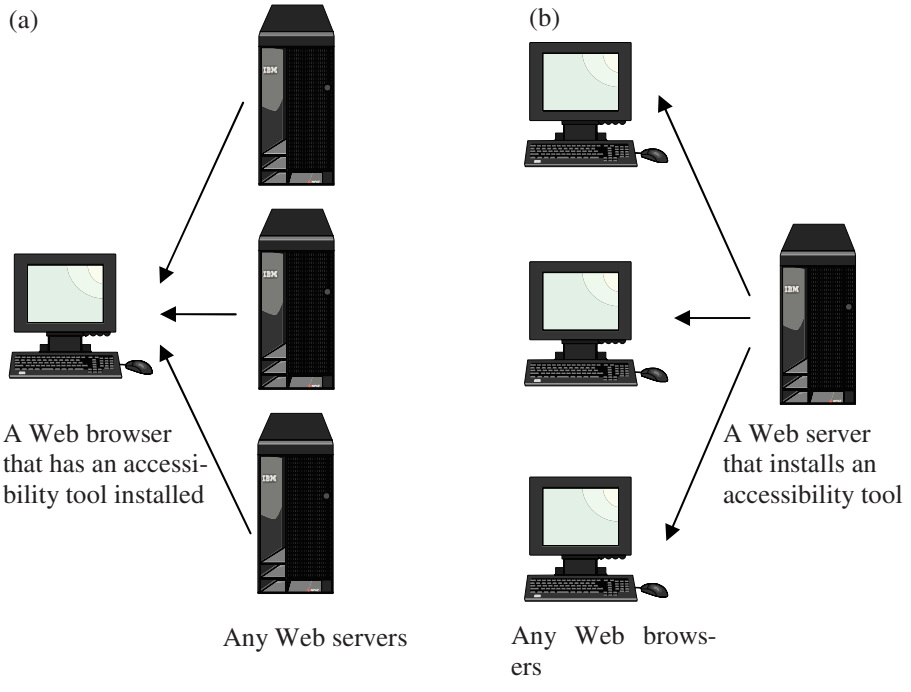


Fig. 1. Two types of accessibility tools: (a) an accessibility tool for a Web browser, and (b) an accessibility tool for a Web server

Fig. 2 shows three methods for a website's owner to improve the accessibility of the website. The first method is by generating personalized webpages on the Web server (a). The text size, the text color, and the background color of the webpages are changed according to each user's preferences. This method is not acceptable for many website owners, because they already have a lot of content in their website and this method usually requires modification of the existing content to parameterize it. The second method is transcoding webpages at a Web intermediary [5, 6, 7] (b). This method does not require modification of the existing content, but lots of computing power is required for the Web intermediary if the number of users is large. The third method involves sending a program to the Web browsers, and the program modifies the appearance of the webpages on the Web Browsers [8, 9, 10] (c). This method also requires no modification of the existing content. In addition, this method is scalable, because the changes of the webpages' appearances are made by the users' machines. This paper describes an approach using the third method.

Most of the proposed solutions [8, 9, 10] to improve websites' accessibility use technologies available only on a specific client environment (such as ActiveX for Internet Explorer). With such solutions, users have to use a specific operating system and a specific Web browser. Our work addresses this problem.

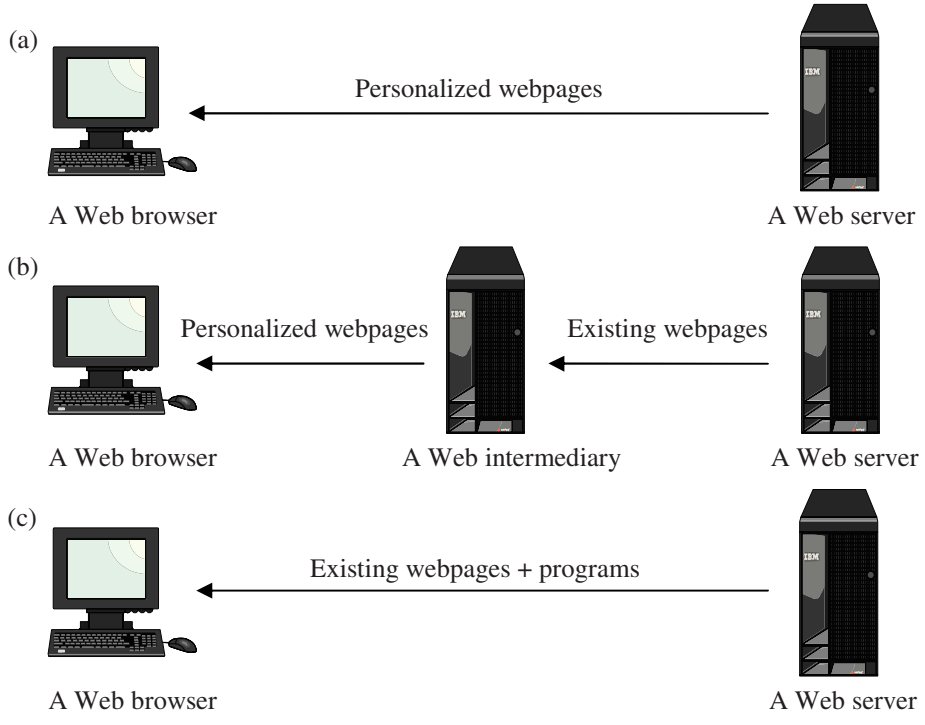


Fig. 2. Three types of method to improve website accessibility: (a) generating personalized webpages on a Web server, (b) transcoding webpages by a Web intermediary, and (c) sending a program to a Web browser

The remainder of this paper is organized in the following way: Section 2 presents the basic methods to improve accessibility for an existing website using JavaScript™. Section 3 suggests ways of enhancing the method for related websites in multiple domains. Section 4 concludes the paper and describes our future work.

2 Improving Accessibility Using JavaScript

Our proposed method uses JavaScript [11] to add new functions to improve accessibility. JavaScript is supported in many Web browsers and such Web browsers run on many operating systems, so users can access these accessible websites from heterogeneous client environments. In addition, users do not have to install a program when using JavaScript. JavaScript programs are executed without special installation.

A webpage loaded in a Web browser is represented by a tree structure called a Document Object Model (DOM) [12]. JavaScript can handle events generated by a DOM, and can modify the DOM to change the webpage's appearance. For example, JavaScript can highlight a text element under the mouse pointer, and can show the text in another frame with larger text. Fig. 3 shows an example of enlarged text.

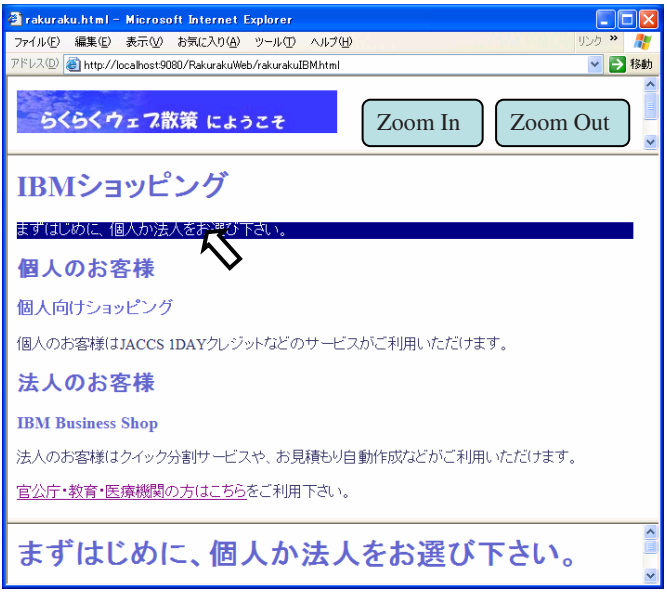


Fig. 3. Example of enlarged text

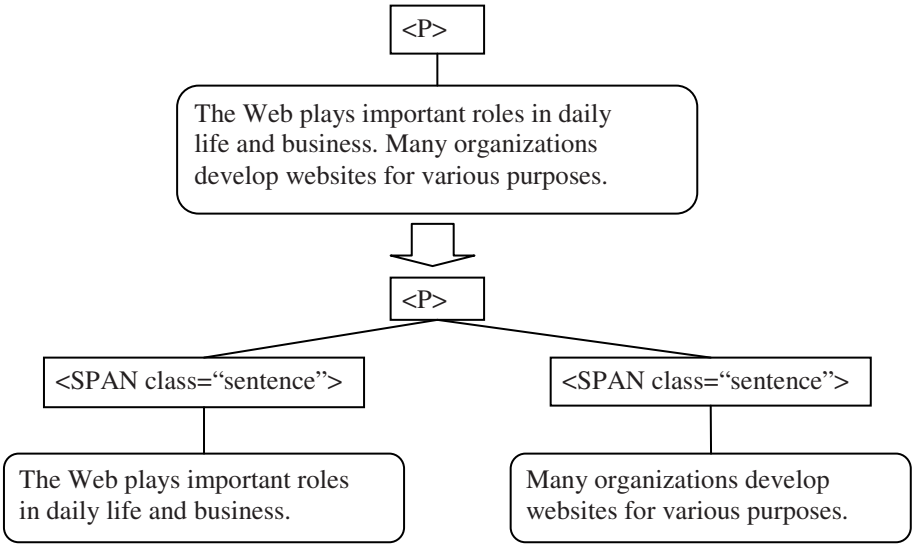


Fig. 4. Modification of a DOM

Our method uses several frames: one frame (the content frame) is used for displaying existing webpages and the other frames are used for additional functions such as displaying buttons and showing enlarged text. When an existing webpage is loaded in the content frame, an "onload" event handler for the content frame is called. (Although the

"onload" event handler for the "FRAME" element is not actually defined in HTML specification [13], it is enhanced in both of the major Web browsers Internet Explorer and Mozilla.) The handler modifies the appearance of the webpage and sets event handlers for "mouseover" and "mouseout" events on the page. By setting up the event handlers, we can manipulate the webpage's DOM when events are generated.

The "mouseover" event occurs when the pointing device is moved onto an HTML element, and the "mouseout" event occurs when the pointing device is moved so it is no longer over an HTML element. We can highlight the pointing HTML element, and show the enlarged text in the other frame by the event handlers.

It is sometimes necessary to modify the DOM in the "onload" event handler to access fine-grained mouse events. For example, a "P" element can be divided into several "SPAN" elements to get mouse events for the individual sentences (Fig. 4).

When a link in a webpage (in the content frame) is followed to another webpage in the same domain, the "onload" event handler is executed and the accessibility functions work fine. However if it is followed to another webpage in another domain, the "onload" event handler cannot be executed and the accessibility functions do not work. Script programs cannot access a DOM loaded from another domain. This is called the Same Origin Policy [14].

3 Enhancement for Multiple Domains

We enhanced our method so that the added functions for improving accessibility can be used without requiring the users to make special efforts between related websites

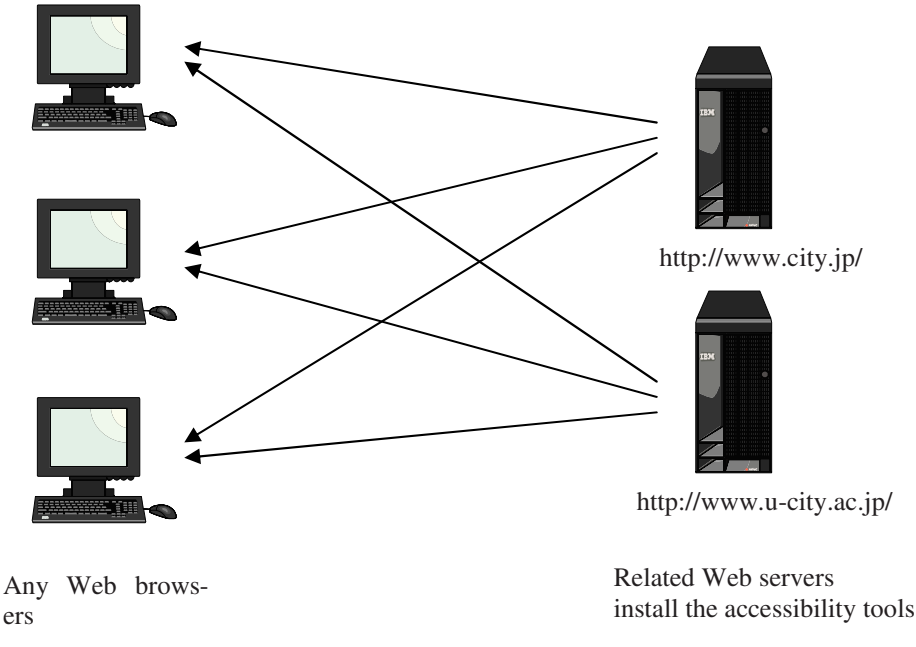


Fig. 5. Related Web servers install the accessibility tools

in multiple domains. Suppose a website's owner want to improve the accessibility for two related websites such as http://www.city.jp/ and http://www.u-city.ac.jp/ (Fig. 5). And some of the webpages in the websites are linked to each other. Our goal is to let users navigate from one website to the related website while using the accessibility capabilities without special efforts.

Each website has a dynamic webpage for generating an accessible page with a frameset: http://www.city.jp/EasyWeb and http://www.u-city.ac.jp/EasyWeb. The URL of the content frame is specified as a parameter. It is possible to access the webpage http://www.city.jp/news.html with improved accessibility by specifying http://www.city.jp/EasyWeb?content=news.html.

Following is an example of generating an accessible webpage implemented with JavaServer Pages™. The second frame is a content frame, and the "src" attribute is specified by the "content" parameter. The other frames are used for additional functions.

```
<HTML>
<HEAD>
<%@ page language="java" contentType="text/html" %>
<TITLE>Easy Web</TITLE>
<SCRIPT type="text/javascript" src="easy.js"></SCRIPT>
</HEAD>
<FRAMESET rows="20%,*,30%">
  <FRAME src="controller.html" name="controllerFrame">
  <FRAME src="<%=request.getParameter("content")%>"
    name="contentFrame" onload="init(contentFrame)">
  <FRAME src="magnifier.html" name="magnifierFrame">
</FRAMESET>
</HTML>
```

We added the following two steps to the "onload" event handler for the content frame to support multiple domains. First, we check the "href" attribute of all "A" elements and the "action" attribute of all "FORM" elements in the content frame's page. Next, if the attribute value is specifying a page in a related website, we change the value to be a corresponding URL for improved accessibility, and change the "target" attribute to "_top". "_top" means that loading a new webpage does not occur into the content frame but into the topmost window. When the user navigates to the modified URL, the webpage contains a "FRAMESET" and the webpage in the content frame have the same origin so the "onload" event handler can execute. By using these steps, users can navigate to the pages of the related websites with improved accessibility.

4 Conclusion and Future Work

We developed a prototype system for our proposed method, and confirmed that our method can improve accessibility for existing websites spanning multiple domains. We also confirmed that our method works on several combinations of operating

system and Web browser: Windows™ with Internet Explorer, Windows with Mozilla, Linux™ with Mozilla, and Mac with Mozilla.

The IBM® Tokyo Research Laboratory developed an assistive Web browsing system, Easy Web Browsing [8, 9], in 2002. The system has already been adopted by some IT leaders among local governments, ministries of the central government, and IT companies in Japan and other countries. It enlarges the displayed characters, changes background and character colors, and reads text aloud. The currently released version is implemented with ActiveX, so it works only on the combination of Windows with Internet Explorer. Although installing Easy Web Browsing is not difficult, users do have to install it on their computers.

We plan to enhance Easy Web Browsing by using our new method. Although it is possible to implement most functions required for improving the accessibility of webpages by using JavaScript, it is not possible to implement some functions that way. For example, reading text aloud is not possible with JavaScript. We have to implement such functions by using another technology, and call the functions from JavaScript.

References

1. World Wide Web Consortium, Web Accessibility Initiative (WAI) http://www.w3.org/WAI/
2. Section 508: The Road to Accessibility http://www.section508.gov/
3. Rowan, M., Gregor, P., Sloan, D., Booth, P.: Evaluating Web Resources for Disability Access. In: Proceedings of the 4th ACM Conference on Assistive Technologies, pp. 80–84 (2000)
4. Asakawa, C., Itoh, T.: User Interface of a Home Page Reader. In: Proceedings of the 3rd ACM Conference on Assistive Technologies, pp. 149–156 (1998)
5. Barrett, R., Maglio, P.P.: Intermediaries: An approach to manipulating information streams. IBM System Journal 38(4), 629–641 (1999)
6. Huang, A.W., Sundaresan, N.: A Semantic Transcoding System to Adapt Web Services for Users with Disabilities. In: Proceedings of the 4th ACM Conference on Assistive Technologies, pp. 156–163 (2000)
7. Fairweather, P.G., Hanson, V.L., Detweiler, S.R., Schwerdtfeger, R.S.: From Assistive Technology to a Web Accessibility Service. In: Proceedings of the 5th ACM Conference on Assistive Technologies, pp. 4–8 (2002)
8. Muta, H., Ohko, T., Yoshinaga, H.: An Active-X-based Accessibility Solution For Senior Citizens. In: Proceedings of CSUS's 20th Annual International Conference (2000)
9. IBM Japan, Easy Web Browsing http://www.ibm.com/able/solution_offerings/EasyWebBrowsing.html
10. Hitachi Government & Public Corporation System Engineering, ZoomSight http://www.zoomsight.jp/demo/zsmd/lang/en/html/help/about_help.html
11. ECMA International, ECMAScript Language Specification, Standard ECMA-262, 3rd edn. (December 1999) http://www.ecma-international.org/publications/files/ECMA-ST/Ecma-262.pdf
12. World Wide Web Consortium, Document Object Model (DOM) Level 2 HTML Specification Version 1.0, W3C Recommendation (January 09, 2003) http://www.w3.org/TR/2003/REC-DOM-Level-2-HTML-20030109/

13. World Wide Web Consortium, HTML 4.01 Specification, W3C Recommendation (December 24, 1999) http://www.w3.org/TR/1999/REC-html401-19991224/
14. Ruderman, J.: The Same Origin Policy
http://www.mozilla.org/projects/security/components/same-origin.html

IBM is a registered trademark of International Business Machines Corporation in the United States, other countries, or both.
Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.
Windows is a trademark of Microsoft Corporation in the United States, other countries, or both.
Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Large Scale Web Accessibility Evaluation - A European Perspective

Mikael Snaprud and Agata Sawicka

Faculty of Engineering, Agder University College, NO-4876, Norway
{Mikael.Snaprud,Agata.Sawicka}@hia.no

Abstract. The 2002 eEurope Action Plan advocated adoption of the WAI Content Guidelines in public EU web sites. Many of the European countries have performed web accessibility evaluations. Most of the evaluations are based on the Web Content Accessibility Guidelines from W3C/WAI. Still, a range of different evaluation methodologies and scoring schemes are deployed across the member states. This makes it hard to compare the web accessibility status between the different EU countries. The European Accessibility Internet Observatory aims to addresses this problem. The Observatory will produce automated, large scale web accessibility measurements based on which a range of monthly benchmarks will be produced and published online. In this way, the benchmarks should provide a useful input to the eAccessibility and eInclusion policy making.

Keywords: web accessibility, benchmarking, large scale evaluation, eInclusion, policy development.

1 Introduction

In the recent Riga declaration,[1] the EU Member States and the European Commission have confirmed their commitment to ensuring accessibility of all public web sites by 2010. The Web Content Accessibility Guidelines from W3C/WAI indicate the desired accessibility standards.[2] Most of the European countries have now implemented explicit policies on eAccessibility. Many also conduct evaluations based on the W3C/WAI Web Content Accessibility Guidelines. Still, a range of different evaluation methodologies and scoring schemes are deployed across the member states. This makes it hard to compare the web accessibility status between the different EU countries. Moreover, most evaluations involve manual assessments, hindering frequent assessments and decreasing comparability of the results. A uniform, large scale evaluation/benchmarking apparatus is needed to support a coherent evaluation of progress across EU with regard to the WCAG-compliance of the public web resources. A prototype of such apparatus was developed under the EU's Sixth Framework Programme by the European Internet Accessibility Observatory (EIAO) [1] [2].

[1] http://ec.europa.eu/information_society/events/ict_riga_2006/doc/declaration_riga.pdf
[2] http://www.w3.org/TR/WAI-WEBCONTENT/

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 150–159, 2007.
© Springer-Verlag Berlin Heidelberg 2007

The paper is organized as follows: First, a brief review of the current eAccessibility benchmarking practices in Europe is presented and discussed. Next, the European Accessibility Internet Observatory is described, focusing on the mechanisms developed to support flexible data presentation. Concluding remarks are provided in the final section.

2 Accessibility Benchmarking Across Europe

In accordance with the 2002 eEurope Action Plan, most of the European countries have now a national strategy/policy advocating conformance with W3C WAI/WCAG. Several countries have also introduced their own guidelines or even laws in the area of eAccessibility. See Table 1 for some examples.

Table 1. Overview of accessibility guidelines/ laws in a selection of countries. The URL indicates where the guidelines or the law is described on the web as accessed 2007-02-15.

Country	Guidelines/law	URL
Denmark	Guidelines for Internet publishing	http://www.netsteder.dk/publ/index.html
Finland	JHS 129 Guidelines for designing web services in the public administration	http://www.jhs-suositukset.fi/suomi/jhs129
France	Circulaire du 7 octobre 1999 relative aux sites internet des services et des établissements publics de l'Etat	http://www.admi.net/jo/19991012/PRMX9903708C.html
Germany	Barrierefreie Informationstechnik Verordnung	http://bundesrecht.juris.de/bitv/index.html
The Netherlands	Guidelines for accessibility and sustainability of government web sites	http://webrichtlijnen.overheid.nl/
Ireland	The Irish Guidelines for Web Accessibility	http://accessit.nda.ie/guideline index_1.html
Norway	Quality criteria covering accessibility, usability and usefulness. The guidelines is a Norwegian translation of WCAG/WAI	http://www.norge.no/kvalitet/kvalitet2006/kriterier.asp
Sweden	Vägledning 24-timmars webben 2.0	http://www.verva.se/web/t/Page____1154.aspx
UK	The Guidelines for UK government websites. Those guidelines are also supplemented by a guide for public procurement, Guide to good practice in commissioning accessible websites, PAS 78:2006	www.drc-gb.org/pas http://www.cabinetoffice.gov.uk/e-government/resources/handbook/html/htmlindex.asp

A collection of links to accessibility guidelines on government web sites is available from the site Standards Schmandards.[3] A list of policies related to web accessibility is maintained by W3C.[4]

[3] http://www.standards-schmandards.com/projects/government-guidelines/
[4] http://www.w3.org/WAI/Policy/

To yield tangible results, the policy initiatives and the guidelines should be followed by systematic measurement and monitoring activities. It would also be advantageous if the national progress was measured in such a way as to allow for international comparisons. One of the first comprehensive pan-European eAccessibility assessments was the Top of the Web survey carried out in November 2003.

The 2003 survey has been followed in 2005 with the UK Cabinet Office eAccessibility survey. The assessment was conducted for 436 web sites and covered 43 WCAG checkpoints. 13 WCAG checkpoints were evaluated using automated tools; others were assessed manually. The results were presented using custom percentage scale, with no possibility of acquiring detailed information regarding the scores of particular countries/web sites on specific WCAG checkpoints. The report was published in a PDF document available online. As such, it prohibited custom data drills and sorting.

While providing a useful European overview of the eAccessibility status, the 2005 survey has not yet been followed by similar pan-European assessments. There are several eGovernment surveys carried out on an annual basis.

The eGovernment readiness report from the United Nations[5] covers the "official" national web sites and a selection of ministry level web sites for each country. The 2005 survey included a WCAG level 1 conformance check by using the using the online tool from Watchfire.[6] A list of the counties conforming is presented in the report. The series of Web Based Survey on Electronic Public Services Reports prepared by Cap Gemini for the European Commission has in the 6th[7] and most recent measurement report included accessibility for assessment, although not in the final score. These reports does however not yet include detailed quantifiable information the development accessibility over time.

While there seems to be no continuous pan-European eAccessibility measurement and monitoring initiative, there is a range of interesting initiatives at national level. Some of the most advanced eAccessibility evaluation activities on national level are today carried out by Denmark, Netherlands, Norway and Sweden. On the pan-European level, there has recently been one evaluation effort undertaken, during the UK Presidency. Table 2 provides an overview of 5 national benchmarking initiatives.

In all cases, the measurements are performed by public agencies. Most of the evaluation schemes are quite recent, with only the Danish *Best on Web* and Norwegian Norge.no dating back to 2000 and 2001, respectively. The benchmarks are typically published on a yearly basis, with an opportunity to view past results. The exceptions are the Swedish benchmark that is issued quarterly and the Dutch Overheid.nl Monitor that evaluate at least once a year, and in addition on demand as required by the web site owner or the team responsible for the Monitor.

---

[5] http://unpan1.un.org/intradoc/groups/public/documents/un/unpan021888.pdf

[6] http://webxact.watchfire.com/

[7] http://ec.europa.eu/information_society/eeurope/i2010/docs/benchmarking/online_availability_2006.pdf

Table 2. Comparative review of 5 national eAccessbility benchmarking initiatives

	DK[8]	DK[9]	NL[10]	NO[11]	SE[12]
General information					
Responsible agency	public	public	public	public	public
First evaluation	2006	2000	2005	2001	2005
Frequency	?	yearly	yearly	yearly	quarterly
Past results availability	yes	yes, from 2001	yes	yes, apart from 2002	yes
Languages	Danish	Danish	Dutch	Norwegian	Swedish
Evaluation scope					
URLs, 2006	122	800	524	691	1000
Evaluation areas	accessibility	usability, digital self-service, openness, technical accessibility	accessibility	usability, usefulness, accessibility	accessibility Dublin Core
Accessibility criteria	45 WCAG 8 national guidelines[13]	12 WCAG	14 WCAG	10 WCAG	10 WCAG + Custom
Evaluation method	semi-automated	semi-automated[14]	automated + manual	semi-automated	automated
Data presentation					
Publishing outlet	online open office documents	web	web	web	web / spreadsheet
Dynamic sorting	no	yes	no	yes	yes (spreadsheet)
Regional aggregation	no	no	yes	no	no
Results format	number of failed tests	custom percentage/ point scales	custom percentage scale	custom percentage scale	[yes/no] / No of instances, failed tests
Detailed WCAG test results	no	yes	no	no	partly

The number of URLs covered by the evaluations differs greatly from 122 web sites covered by the oio.dk Danish initiative, to about 1000 of web sites covered by the Swedish evaluation. With the exception of the Swedish initiative and the report issued

[8] Danish portal on public information, http://www.oio.dk/tilgaengelighed/ kortlaegning07
[9] Danish Top of the Web project, http://www.bedstpaanettet.dk
[10] Dutch portal to public information, http://webrichtlijnen.overheid.nl/
[11] Gateway to the public sector in Norway, http://www.norway.no/
[12] Verva -Swedish Administrative Development Agency, http://www.verva.se
[13] The additional national guidelines concern the use of electronic forms and PDF documents.
[14] The manual evaluations are carried out by two independent experts.

by the UK Office Cabinet, all the evaluations cover also other areas beyond the accessibility. With respect to the accessibility assessment, all the evaluations draw on the W3C/WCAG checkpoints, although they differ in their coverage scope.

The evaluations are carried out in different ways, the Dutch and the Swedish initiatives are largely automated, whereas the others are based on semi-automated assessments involving manual expert evaluation supported by the results of the automated W3C validators where appropriate. The UK Cabinet Office details which of the WCAG checkpoints are evaluated exclusively manually and which involve the automated checks. The results of the assessments are in all cases available online, supporting data drilling and sorting to varied degree. None of the reviewed tools supports aggregation across sectors. Only the Dutch monitoring tool supports a dynamic aggregation/disaggregation across regions; the other reports provide the eAccessibility status for individual web site or group of web sites belonging to one particular region. In most cases, a single eAccessibility measure is provided and expressed using a custom percentage scale. With an exception of the Danish *Best on Web* initiative, none of the reports supports disaggregation of the overall score to the level of the specific WCAG checkpoints tests.

The evaluation initiatives briefly outlined above, illustrate a range of interesting initiatives. The potential impact of the measurements could potentially increase if pan-European comparisons where supported.

A comparison of the initiatives show that the main elements of the methods used, from sampling, though guidelines and evaluation to the presentation of results are different among the counties. In particular, the evaluation, there are different variants of national guidelines based on WCAG. Also the schemes to assign an overall score for a web site can include additional aspects of the web site like usefulness or usability in Norway [1], and four additional aspects are included in the score in the Netherlands. To locate good practice and to encourage an element of competition towards a more inclusive Information Society, a unified web evaluation methodology is needed.

3 The EIAO Approach

The WAB (Web Accessibility Benchmarking)[15] cluster is co-funded by the European Commission, and was organised to contribute to harmonisation of the efforts towards an inclusive information society in Europe. The main activity of the cluster is to define a unified web evaluation methodology (UWEM) for conformance to the WCAG1.0 guidelines and ensuring a smooth transition to the announced WCAG2.0 guidelines. The cluster initially consisted of the project EIAO[16], Support-EAM and BenToWeb.

The European Internet Accessibility Observatory (EIAO, www.eiao.net) currently develops a prototype online Observatory to perform large scale automatic accessibility benchmarking based on the UWEM1.0, with a plan to evaluate and

[15] http://www.wabcluster.org/

[16] The project is co-funded by the European Commission DG Information Society and Media, under the contract IST-004526.

present results from 10 000 public European web sites monthly. The first results from the large scale evaluation is due for public release in June 2007. All the software developed for the Observatory is developed and released as Open Source [3][4] to encourage review, facilitate participatory design [5][6], and pave the way for further improvements. This design decision has proved particularly useful to invite new developers to the team and reuse the solutions provided by other projects.

The results from the prototype implementation indicate that adapting and integrating existing tools rather than building a system from scratch has been very suitable for the task. This choice has also, as implemented by the highly talented development team, resulted in an extensible architecture for large scale web accessibility evaluation [7][8].

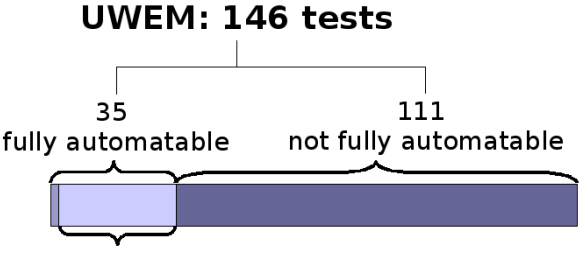


Fig. 1. Coverage of UWEM tests in the EIAO demonstrator

UWEM contains in total 146 tests. 35 of these tests can be fully automated (see Figure 1), and 31 of these are implemented by the EIAO WAMs (covering 19 WCAG checkpoints). In addition, results are included from 2 UWEM tests provided by the BenToWeb project.[17]

Only automated tests are viable to use for large scale testing. However, to cover all tests manual and labour intensive evaluation is needed. The large scale evaluation is expected to be useful to identify and disseminate good practice from a larger sample of web sites. A combination of both could be thus be used to cater for a low cost and frequent evaluation to be supplemented by more in depth manual evaluation of particularly interesting web sites.

3.1 Presentation of the Collected Evaluation Results

The reporting capacity of the Observatory is designed to allow for easy generation of benchmarking lists, as well as to support more advanced functionality for drilling down into details on a given web site. The reports have been designed primarily to support policy makers in their work [9] .They should be accessible through a flexible and yet intuitive user interface that facilitates analysis of the data in the data warehouse from different perspectives [10][11].

The most aggregated score shows the overall barrier probability for all the web sites covered in Europe. From this aggregated level the user can drill down through

[17] http://bentoweb.org/

the county level, and further down to a county and even onwards to a given web site and the barriers detected on this particular web site. Figure 2 illustrates the drill-down and roll-up navigation of the reports. Please note that the data in the following figures are included for demonstration purposes only.

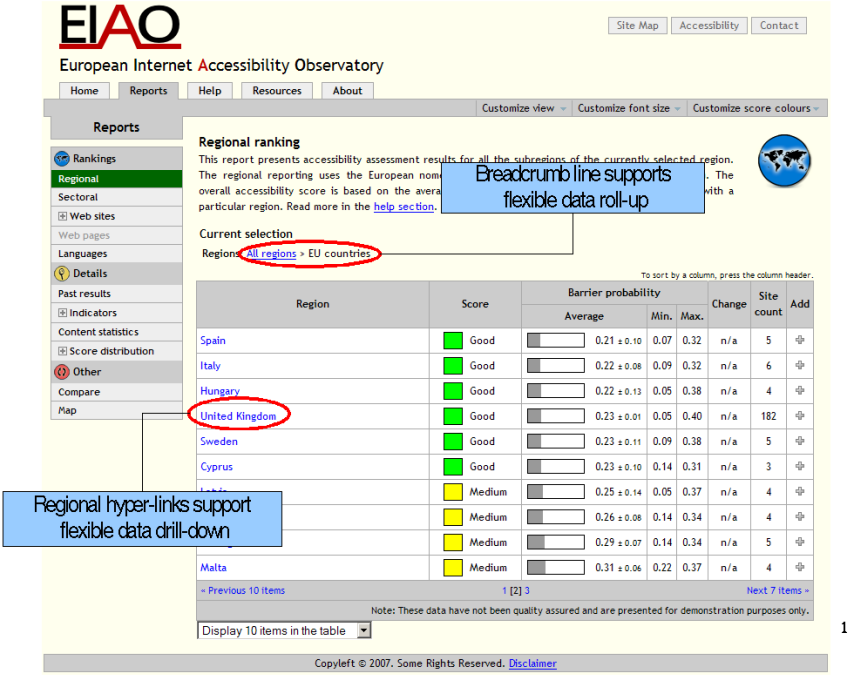


Fig. 2. The EIAO reports support drill-down and roll-up for navigation in the data warehouse. Please note that the data in the figure is included for demonstration purposes only

Aggregations can be performed also on other items like the detected barriers among on all the sites in Europe or in a given country, For each report the columns can be sorted as needed allowing for different perspectives of the data. see Figure 3.

The barriers shown in Figure 3 are corresponding to the set of UWEM tests that can be automated according to UWEM 1.0. A similar view can be shown also for the barriers discovered for a given web site. Since a site is represented by a limited number of web pages from the evaluated web site, the detected barriers will only be a representative selection, not all barriers that could be detected automatically on all pages of the web site. The selection of web pages is carried out according to a random walk algorithm.

An other form of aggregations is illustrated in Figure 4, where the distribution of web sites among scores as defined in UWEM 1.0 is depicted. In addition figure 4 shows a geographical to give a further aggregated view of the results.

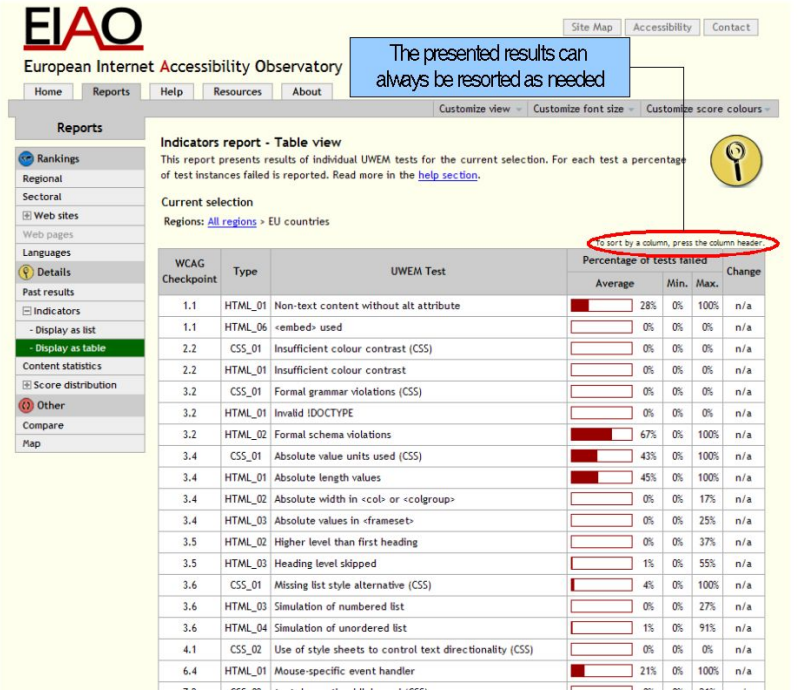


Fig. 3. Aggregation on detected barriers among all web sites evaluated across Europe. Please note that the data in the figure is included for demonstration purposes only.

In addition to the current status of the accessibility barriers on public web sites, the reports have been designed to show the development over time. In this way, policy makers may be able to monitor the longer term impact of the implemented policies.

The approach to visualise the data could also be used present other kinds of data. For example the accessibility evaluation could be extended with additional indicators to cover other aspects of eGovernment development such as update frequency of web content or to what extent citizens are requested to fill in the same information in different government forms.

4 Concluding Remarks

Several national benchmarking initiatives are already producing very interesting results and publish them on a regular basis. They are all based on WCAG1.0 and have several similar features. However, due to differences in sampling, in the interpretation of WCAG1.0, and in the representation of the evaluation results, the comparison of results among different measurements is hardly feasible.

Wide use of a common evaluation methodology, based on WCAG, would facilitate international comparisons and enhance the spread of good practice.

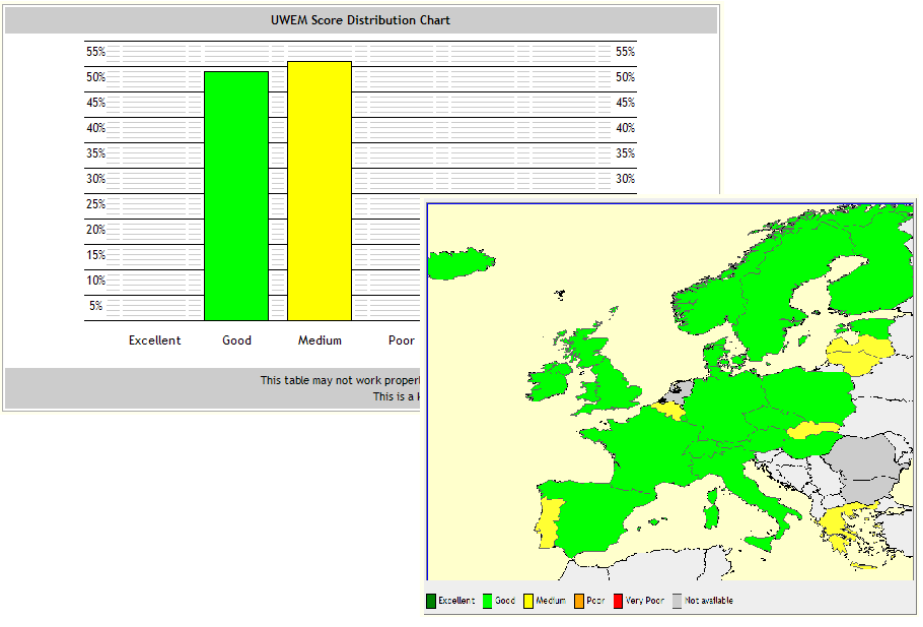


Fig. 4. Distribution of web sites among the barrier probability scores as defined in UWEM, shown as bar chart and as geographical map. Please note that the data in the figure is included for demonstration purposes only.

In this way, the unified web evaluation methodology (UWEM) based on WCAG from W3C, can fuel the development towards a more Inclusive European Information Society. A demonstrator for UWEM 1.0 has been implemented as an open source application by the EIAO project. The the first large scale evaluation results from 10.000 European public web sites are now available as online reports.

Acknowledgements. Please note that the results presented in this paper are built on the achievements of the open source community, and in is the result of a team effort with substantial contributions from 10 organisations. You find more details on the EIAO project team, the organisations, and their contributions on the project web site: http://www.eiao.net.

References

1. Aslaksen, F., Fardal, F., Snaprud, M.: The Role of Benchmarking in Concerted Actions to Increase Accessibility, 10th International Conference on Computers Helping People with Special Needs, ICCHP 2006, 12-14 July 2006, Linz. In: Miesenberger, K., Klaus, J., Zagler, W., Karshmer, A.I. (eds.) ICCHP 2006. LNCS, vol. 4061, pp. 258–262. Springer, Heidelberg (2006)
2. Craven, J., Snaprud, M.: Positive Steps to Assessing, Monitoring and Comparing the Accessibility of Public Websites: Introducing the European Internet Accessibility Observatory, FreePint (August 24, 2006) pp. 18–22 (2006)

3. Perens, B.: The open source definition. In: DiBona, C., Ockman, S., Stone, M. (eds.) Open Sources: Voices from the Open Source Revolution. O'Reilly and Associates, Cambridge, Massachusetts (1999)
4. Duque, R., Collins, M., Abbate, J., Azambuja, C.C., Snaprud, M.: History of Information Technology. In: Shrum, W., Benson, K.R., Bijker, W.E., Brunnstein, K. (eds.) Past, Present and Future of Research in the Information Society, Springer, Heidelberg (2007)
5. Nygaard, K., Bergo, T.O.: Planlegging, styring og databehandling. Grunnbok for fagbevegelsen (Planning, management and data processing. Handbook for the labour movement) vol. I. Oslo: Tiden norsk forlag (1973)
6. Mohaghegi, P., Snaprud, M., Pillai, A., Ulltveit-Moe, N.: Including Users and External Contributors in an European Research Project, Digital Inclusion and Open Source October 20 -21, 2005 Oslo, Norway (2005)
7. Snaprud, M.H., Ulltveit-Moe, N., Pillai, A.B., Olsen, M.G.: A Proposed Architecture for Large Scale Web Accessibility Assessment. In: Miesenberger, K., et al. (eds.) ICCHP 2006. LNCS, vol. 4061, pp. 234–241. Springer, Heidelberg (2006)
8. Snaprud, M.: Early results from automatic accessibility benchmarking of public European web sites from the European Internet Accessibility Observatory (EIAO). In: Proc. of Design for All, DFA, September 13-15 2006, Rovaniemi, Finland (2006)
9. Snaprud, M., Sawicka, A.: The European Internet Accessibility Observatory project, UNPAN, Expert Group Meeting on E-Participation and E-Government, Budapest, Hungary July 27-28 2006 proceedings yet (to appear) (2006) paper available at http://unpan1.un.org/intradoc/groups/public/documents/UN/UNPAN023676.pdf
10. Sawicka, A., Snaprud, M., Unander, E.H.: Online Presentation of Large Volume Web Accessibility Data Within EIAO. In: The second Workshop on Web Accessibility and Metamodelling (WWAM 2006) June 13-15 2006 Agder University College, Grimstad, Norway (2006)
11. Sawicka, A., Snaprud, M.: eAccessibility Data for eInclusion Policy Making: Reports from a Large Scale Internet Observatory. In: Proc. of Design for All, DFA, September 13-15, 2006, Rovaniemi, Finland (2006)

The BenToWeb XHTML 1.0 Test Suite for the Web Content Accessibility Guidelines 2.0 - Last Call Working Draft

Christophe Strobbe[1], Jan Engelen[1], Johannes Koch[2], Carlos Velasco[2], Evangelos Vlachogiannis[3], and Daniela Ortner[4]

[1] Katholieke Universiteit Leuven, Kasteelpark Arenberg 10, B-3001 Heverlee-Leuven, Belgium
{Christophe.Strobbe,Jan.Engelen}@esat.kuleuven.be
[2] Fraunhofer-Institut für Angewandte Informationstechnik (FIT), Schloss Birlinghoven, D-53757 Sankt Augustin, Germany
{Johannes.Koch,Carlos.Velasco}@fit.fraunhofer.de
[3] University of the Aegean, Voulgaraktonou 30, GR11472 Exarchia, Athens, Greece
evlach@aegean.gr
[4] University of Linz "integriert studieren - integrated study" (i3s3), Altenbergerstrasse 69, A-4040 Linz, Austria
Daniela.Ortner@jku.at

Abstract. This paper presents the work carried out under the umbrella of the EU-funded project BenToWeb to develop a complete XHTML 1.0 test suite for the W3C's Web Content Accessibility Guidelines 2.0. Initial work covered the June 2005 working draft, which was subsequently updated to the April 2006 working draft ("last call"). At the time of writing, a thorough evaluation, involving end users, is being carried out.

Keywords: test suite, accessibility, Web Content Accessibility Guidelines (WCAG), evaluation and repair tools (ERT), HTML, XHTML, CSS, Web Accessibility Initiative (WAI), World Wide Web Consortium (W3C), Last Call Working Draft.

1 Introduction

The EC-funded project BenToWeb (Benchmarking Tools and Methods for the Web[1]) has a goal to develop test suites for the forthcoming Web Content Accessibility Guidelines 2.0. The development of these test suites serves a dual purpose. First, they support the WAI Working Groups in the development of support documents, such as technology-specific techniques, for WCAG 2.0. The test files prove whether techniques can be implemented, and help find out ambiguities and loopholes in WCAG documents. Second, the test suite can be used to benchmark accessibility evaluation and repair tools (ERT). Failing test files can also be used to test

[1] http://www.bentoweb.org/

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 160–166, 2007.
© Springer-Verlag Berlin Heidelberg 2007

conformance to the Authoring Tools Accessibility Guidelines (ATAG)[2], for example to test if a tool's user interface allows authors to avoid flashing in a WYSIWYG editing view.

WCAG 2.0 currently has general techniques, HTML techniques, CSS techniques, client-side scripting techniques and server-side techniques. BenToWeb will not cover all of these technologies exhaustively, but at least XHTML 1.0 + CSS 2.0. In 2005, BenToWeb created a test suite for the June 2005 working draft of WCAG 2.0. After the WCAG Working Group published the Last Call Working Draft of WCAG 2.0[3] in April 2006, this test suite was updated to the newest draft.

2 Testing and Comparing Evaluation and Repair Tools

Since the publication of Web Content Accessibility Guidelines (WCAG) 1.0, there have been many efforts to automate the evaluation of web pages or even complete websites against these guidelines. There is great variability in the strengths and features of these tools, and comparing them is a time-consuming task. Several researchers have already reported on this type of work, including Brajnik [1] and Melody Ivory [3,4,5]. The problem with these comparisons is that they usually rely on web pages or on samples from websites that are not available for other researchers to repeat the study with other tools. The outcome of the research is partially based on an unknown input. A publicly available suite of test cases that have been individually described and validated, would eliminate this the problem of repeatability. For each test case, it should be clear which accessibility guidelines are violated (or that specific guidelines are not violated), and how often they are violated. This documentation can then be compared with the report of an evaluation tool that has "checked" the test suite, in order to find out what guidelines are covered by the product, what guidelines trigger false positives, false negatives, etcetera. A publicly available test suite thus enables comparisons or benchmarking of products, and comparisons of the same product over time.

3 Other Accessibility Test Suites

The World Wide Web Consortium's Quality Assurance Activity[4] has a goal to ensure that its deliverables – W3C Recommendations – are implemented correctly. The development of test suites can contribute to this goal. In the area of web accessibility, both inside and outside the Web Accessibility Initiative (WAI), there have been several efforts to build test suites. The scope, size and status of these test suites vary greatly: in the beginning of 2005, it was not clear if any of these could be considered as finalized [6].

There are test suites for the User Agent Accessibility Guidelines 1.0[5], for the Web Content Accessibility Guidelines 2.0[6], and for several software products, such as

[2] http://www.w3.org/TR/ATAG20/
[3] http://www.w3.org/TR/2006/WD-WCAG20-20060427/
[4] http://www.w3.org/QA/
[5] For example at http://www.w3.org/WAI/UA/TS/html401/.
[6] HTML Test Suite for WCAG 2.0: http://www.w3.org/WAI/GL/WCAG20/tests/.

screen readers and the Mozilla browser. The HTML Test Suite for WCAG 2.0 comes closest to the type of test suite developed by BenToWeb. The test suite is no longer being updated, but the WCAG Working Group has integrated tests into the techniques document for WCAG 2.0[7]. In 2006, the Evaluation and Repair Tools Working Group (ERT WG) and the WCAG Working Group also set up a joint "Test Samples Development Task Force", with the objective to develop test samples for WCAG 2.0 Techniques[8].

4 Structure of the Test Suite

In this context, a test suite is not a set of tests that can be used to validate web content, but a set of test files with accompanying metadata both for human and machine consumption. The test suite is a collection of "test cases", where a test case consists of one or more XHTML files that implement or fail a requirement specified by a WCAG 2.0 success criterion, and an accompanying metadata file. The metadata are recorded in an XML format specially created for this purpose: Test Case Description Language (TCDL) [8]. The metadata include a short title and a description of the test file or files, a statement on whether the test files pass or fail the success criterion, the location of the issue if the test files fail, and guidance on what is necessary to validate the test case, for example scenarios for end-user evaluation.

For each WCAG 2.0 success criterion, at least two test cases need to be created: at least one that fails and at least one that passes the success criterion. This is because there needs to be at least one test for a false positive, and at least one test for a false negative, respectively. When the test suite is complete and validated, running the test files through an accessibility evaluation tool should then provide data on the completeness of the tool's coverage of WCAG 2.0 and whether it generates false positives and negatives.

BenToWeb distinguishes between several types of test cases. "Atomic test cases" address only one success criterion and use only a single XHTML file (supporting files such as images, client-side scripts or style sheets do not count in this context; the XHTML file can either be static or generated with JavaServer Pages (JSP)). However, some accessibility requirements apply to sets of web pages instead of pages in isolation: the Last Call Working Draft of WCAG 2.0 contains success criteria about consistency of navigational mechanisms such as site navigation (SC 3.2.3 and 3.2.4), bypassing repeated content (SC 2.4.1), finding information in a website (SC 2.4.2) and information about a user's location in a website (SC 2.4.7). Test cases for these success criteria use multiple XHTML files and are called "compound test cases". Ideally, a test suite for benchmarking evaluation software should also contain test cases for combinations success criteria. In BenToWeb, these would be called "complex test cases". The Test Case Description Language supports the description of such test cases: it is possible to reference multiple success criteria, and to state for each whether the test case passes or fails. Previous accessibility test suites do not appear to contain this type of test cases [6].

[7] Techniques for WCAG 2.0: http://www.w3.org/TR/WCAG20-TECHS/.
[8] WCAG 2.0 Test Samples: http://www.w3.org/WAI/ER/tests/.

5 The Development and Evaluation Process

The development process requires that each test case moves through several steps before it is finally accepted in the test suite. Each test case starts out as a draft and goes through two levels of evaluation. First, it is reviewed by an accessibility or HCI expert. If any issues are found, the test case is sent back to the test case author. These issues may be editorial or content-related. Sometimes they relate to the interpretation of a success criterion. In this phase, the evaluator can also decide that the test case would benefit from end-user evaluation and set up scenarios that match a certain user profile (disabilities, experience with user agents, assistive technologies).

After the author and the first evaluator have solved the issues, the test case is ready for the second level of evaluation. The second evaluator can accept the test case for end-user evaluation, send it back to the test case author, or recommend that the test case be included in or rejected from the test suite. Test cases that contain scenarios for end-user evaluation are loaded into a test case evaluation framework (described by Herramhof (2006) [2]). The framework matches scenarios with user profiles and saves the users' input for later analysis. After evaluation and when all data are definitive, the test case is finally "accepted" into the test suite.

6 Current Status of the Test Suite

6.1 Size of the Test Suite

The first version of the test suite covered the 30 June 2005 Working Draft of WCAG 2.0. It was made up of 477 test cases, each consisting of one metadata file (in TCDL) and one or more content files. The test suite contained over 520 XHTML files (or JSP files that generated XHTML)[9], which often use supporting files, such as JavaScript, CSS, GIF, JPEG, WMV (audio/video), WMA, WAV, MP3 and Java applets.

At the time of writing, the second version of the test suite contains over 530 test cases for the 56 success criteria in the Last Call Working Draft. These test cases contain over 600 XHTML test files (or JSP files that generate XHTML), and a smaller number of supporting files (JavaScript, CSS, images, audio, video and applets). Both versions of the test suite are publicly available[10].

6.2 Interpretation of WCAG Success Criteria

The second version of the test suite covers the 27 April 2006 Working Draft ("last call") of WCAG 2.0. This means that the working draft published on 23 November 2005 was skipped. In the June 2005 working draft, some success criteria were unfinished or open to interpretation. This led to discussions during the development and evaluation of the BenToWeb test suite, and to feedback to the WCAG Working Group. We refer to the conference paper on the first version of the test suite for

[9] The numbers cited here are slightly lower than in a previous conference paper (Strobbe (2006) [7]: 481 test cases, over 530 XHTML files) because that paper was written when the evaluation of the test suite was not yet finished.

[10] http://www.bentoweb.org/ts

examples [7]. Between June 2005 and April 2006, the WCAG Working Group solved many issues and published additional supporting documents: "Understanding WCAG 2.0"[11] and "Techniques for WCAG 2.0"[12] (later supplemented with the "WCAG 2.0 Quick Reference"[13]). As a result, there were fewer interpretation issues in BenToWeb; they were also easier to solve internally and were about specific details rather than a success criterion as a whole. Nevertheless, if after internal consultation, project participants were not fully confident about the interpretation of a success criterion, the issue was fed back to the WCAG Working Group.

6.3 Comprehensiveness of the Test Suite

The publication of the additional supporting documents by the WCAG Working Group also helps BenToWeb to develop a more comprehensive test suite. When the first test suite was developed, the WCAG Working Group had only defined a relatively small set of HTML techniques for WCAG 2.0. Test case authors were free to draw on any documentation of techniques or failures they could find, regardless whether the source was WCAG or not. When the WCAG Working Group published the Last Call Working Draft of WCAG 2.0, it also published "Techniques for WCAG 2.0", containing both techniques and common failures related to accessibility. BenToWeb test cases can now also be mapped to these techniques and failures, which enables a gap analysis. Gap analyses against the HTML and CSS specifications will also be conducted. Evaluating the comprehensiveness of the test suite and searching for techniques and failures not documented by the WCAG Working Group are ongoing tasks.

6.4 Variability in the Test Suite

In the first version of the test suite, some success criteria had only two test cases, while others had more than thirty. This variability is also present in the second version and seems to be inherent in any XHTML test suite for WCAG 2.0. The variability in the number of test cases per success criterion is often related to the number of XHTML elements or attributes that can be used to pass or fail a success criterion. For example, success criterion 3.1.1 requires that "[t]he primary natural language or languages of the Web unit can be programmatically determined." The primary language is usually declared by means of the `lang` and/or the `xml:lang` attribute on the element that contains the complete document (the `html` element). Declaring the primary language on the `body` element may also be acceptable, but this exhausts all the options. There are few techniques to declare the primary language, and few ways to fail this requirement. By contrast, success criterion 1.4.1 requires a luminosity contrast ratio of 5:1 between foreground and background colours. However, colours can be set on basically any element that can appear in the body of a document, and these elements can be nested in many different ways. Success criterion 1.3.1 is an even better example. It requires that "[i]nformation and relationships

[11] http://www.w3.org/TR/UNDERSTANDING-WCAG20/
[12] http://www.w3.org/TR/WCAG20-TECHS/
[13] http://www.w3.org/WAI/WCAG20/quickref/

conveyed through presentation can be programmatically determined, and notification of changes to these is available to user agents, including assistive technologies." This covers the use of heading elements to express the structure of a document, connections between form fields and their labels, semantic markup versus abuse of presentational markup, etcetera. There are many ways to fail this success criterion, each with dozens of variations. For this reason, the variability in the number of test cases per success criterion can be said to be inherent in the test suite.

7 Ongoing and Future Work

The current draft of WCAG 2.0 (the Last Call Working Draft) is not a final document. BenToWeb will update the test suite to the draft that will become available in early 2007. This means that the mapping of the test cases to success criteria will need to be updated. It also means that test cases will need to be created for success criteria that may have been added after the Last Call Working Draft. Moreover, some existing test cases may need to be reviewed because the related success criteria have changed. Other test cases may disappear if the related success criteria have been removed and do not map to other success criteria. In addition to this, more time will be devoted to evaluating the comprehensiveness of the test suite.

BenToWeb participants are also involved in the Test Samples Development Task Force (TSD TF), a joint task force set up by ERT Working Group and the WCAG Working Group with the objective to review the WCAG 2.0 techniques and failures and to develop test samples for them. This task force uses a subset of TCDL for its metadata[14] and the contribution of BenToWeb test cases will facilitate the uptake of BenToWeb's test suite efforts in WAI.

Acknowlegements. This work has been undertaken in the framework of the project BenToWeb — IST-2-004275-STP — funded by the IST Programme of the European Commission.

References

1. Brajnik, G.: Comparing accessibility evaluation tools: a method for tool effectiveness. Univ. Access Inf. Soc. 3, 252–263 (2004) DOI:10.1007/s10209-004-0105-y
2. Herramhof, S., Petrie, H., Strobbe, C., Vlachogiannis, E., Weimann, K., Weber, G., Velasco, C A: Test Case Management Tools for Accessibility Testing. In: Miesenberger, K., et al. (eds.) ICCHP 2006. LNCS, vol. 4061, pp. 215–222. Springer, Heidelberg (2006)
3. Ivory, M.Y, Sinh, R.R, Hearst, M.A: Empirically validated web page design metrics. In: Proceedings of the Conference on Human Factors in Computing Systems (Seattle, WA, March), pp. 53–60. ACM Press, New York (2001)
4. Ivory, M Y, Chevalier, A.: A Study of Automated Web Site Evaluation Tools. Technical Report UW-CSE-02-10-01. University of Washington, Department of Computer Science and Engineering (2002)

[14] WCAG 2.0 Test Samples Metadata: http://www.w3.org/WAI/ER/tests/usingTCDL.

5. Ivory-Ndiaye, M.Y: An Empirical Approach to Automated Web Site Evaluation. Journal of Digital Information Management 1(2), 75–102 (2003)
6. Strobbe, C.: Test-suites State of the Art and Quality Assurance Methods for W3C Recommendations. BenToWeb deliverable D 4.1. (2005) Available at http://www.bentoweb.org/html/BenToWeb_D4.1.html
7. Strobbe, C., Herramhof, S., Vlachogiannis, E., Koch, J., Velasco, C A: The BenToWeb XHTML Test Suite for the Web Content Accessibility Guidelines 2. In: Miesenberger, K., et al. (eds.) ICCHP 2006. LNCS, vol. 4061, pp. 172–175. Springer, Heidelberg (2006)
8. Strobbe, C., Herramhof, S., Vlachogiannis, E., Velasco, C A: Test Case Description Language (TCDL): Test Case Metadata for Conformance Evaluation. In: Miesenberger, K., et al. (eds.) ICCHP 2006. LNCS, vol. 4061, pp. 164–171. Springer, Heidelberg (2006)

DHTML Accessibility Checking Based on Static JavaScript Analysis

Takaaki Tateishi, Hisashi Miyashita, Tabuchi Naoshi, Shin Saito, and Kouichi Ono

Tokyo Research Laboratory, IBM Research
1623-14 Shimotsuruma, Yamato-shi, Kanagawa-ken, 242-8502 Japan
{tate,himi,tabee,shinsa,onono}@jp.ibm.com

Abstract. DHTML accessibility is being standardized by W3C, which provides metadata for UI widgets commonly implemented by HMTL and JavaScript. However it is difficult to check that webpages always have correct metadata according to the standards of DHTML accessibility since UI widgets can be updated by JavaScript programs. Thus we propose a technique for checking accessibility of UI widgets. In this check, we use static program analysis techniques so that we can check accessibility without executing a program. In addition, we developed a prototype system based on the proposed technique and applied it to a simple DHTML application.

Keywords: DHTML accessibility, JavaScript, Static program analysis.

1 Introduction

Rich internet applications based on DHTML are spreading globally. One key component is a client-side scripting language, JavaScript in the most common case. Such a JavaScript program can operate directly on the internal tree representation of a webpage called a DOM (Document Object Model) instance, and contributes to enhance both the functionality and usability of dynamic Web applications. One of the popular methods to enhance the functionality and usability is to create UI widgets such as checkboxes and tab panels. Such UI widgets are represented by a DOM instance and JavaScript programs called event handlers. However a lot of UI widgets are inaccessible since they don't provide information about what is represented by DOM instances, and thus a screen reader cannot identify them as UI widgets. Consequently people with visual disabilities cannot understand what UI widgets are on a webpage.

DHTML accessibility [13, 14, 15] is being standardized by W3C to address such situations, and some of these features are supported by Firefox [5]. The key concept of the standards is a set of accessibility metadata, called accessibility roles and states, for identifying classes of UI widgets such as Checkbox and those valid accessibility states such as "checked" or "unchecked". It also supplies rules for state changes to specify how UI widgets should behave. Assistive technologies such as screen readers can provide useful information for people with visual disabilities by interpreting the accessibility metadata. Thus proper accessibility metadata should be a part of HTML

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 167–176, 2007.
Springer-Verlag Berlin Heidelberg 2007

documents so that people with visual disabilities can understand the structures of the UI widgets and their states.

For the development of DHTML applications, developers should write event handlers for UI widgets that operate on DOM instances according to the DHTML accessibility standards. One of the methods to check accessibility metadata is validation of an HTML document against the standards, but this is not adequate, since we actually need to check the DOM instance each time it is updated by the program, and also check the behavior of the UI widgets. In addition, testing is a popular technique to verify the behavior of software applications as well as dynamic Web applications, but it is hard to check all of the possible DOM instances updated by the program.

In this paper, we propose a novel technique to check that an UI widget behaves according to the DHTML accessibility standards by using a static program analysis technique called DOM analysis [7] to assist developers in writing accessible DHTML applications. Static program analysis involves techniques for predicting values or approximations to the set of possible results that will arise at run-time without running a program. A number of approaches to static program analyses are being studied for different purposes. We also developed a DOM analysis for the purpose of predicting the set of DOM instances updated by a JavaScript program.

With this technique, we formally define a constraint on the accessibility metadata and state changes, called an accessibility rule, for each UI widget based on the DHTML accessibility standards. Such an accessibility rule is described with a state transition diagram and schemata. A state transition diagram consists of states and state transitions, and specifies how the accessibility states are changed. Each schema is associated with a state of the state transition diagram and defines the validity of the roles and states such as the valid states for the roles and the parent-child relationships between roles.

This paper is organized as follows: Section 2 describes the accessibility roles and states with a simple example, and then formalizes accessibility of UI widgets for DHTML applications. In Section 3 we explain how to check the accessibility of UI widgets against the accessibility rules. In Section 4 we comment on our experiments with a prototype system. In Section 5 we briefly describe some related techniques and tools for checking DHTML accessibility. In Section 6 we conclude this paper.

2 DHTML Accessibility and Its Formalization

In this section, we first explain the accessibility roles and states and then formally define the accessibility of an UI widget using a state transition diagram and schemata.

2.1 UI Widgets and Accessibility Metadata

Accessibility roles and states are added to an HTML document as XML attributes for identifying classes of UI widgets and their valid accessibility states. By interpreting the accessibility roles and state, assistive tools such as a screen reader can recognize the HTML element as an UI widget having a valid accessibility state.

Figure 1 is a HTML fragment representing a checkbox. The appearance of the checkbox is defined by the class attribute specifying a style sheet. The x2 and state are namespaces, defined in the drafts of W3C standards [14, 15], for accessibility roles and states respectively.

```
<span class="checkbox" id="chbox1" x2:role="role:checkbox"
      state:checked="true" tabindex="0"
      onkeydown="return checkBoxEvent(event);"
      onclick="return checkBoxEvent(event);" >
a checkbox sample</span>
```

Fig. 1. A HTML fragment representing a checkbox widget

In addition, behavior of the checkbox is implemented by a JavaScript program specified by the onkeydown attribute and the onclick attribute as shown in Figure 2. The program reacts to user events of mouse click and key down, and changes the accessibility state of the checkbox with the setAttributeNS function.

```
function checkBoxEvent(event){
  if ((event.type == "click" && event.button == 0) ||
       (event.type == "keydown" && event.keyCode == 32)) {
    var checkbox = event.target;
    if (checkbox.getAttributeNS("http://www.w3.org/2005/07/aaa", "checked")
        == "true") {
      checkbox.setAttributeNS("http://www.w3.org/2005/07/aaa" "checked",
                              ”checked”,”false”);
    } else {
      checkbox.setAttributeNS("http://www.w3.org/2005/07/aaa",
                              "checked",”true”);
    }
    return false;
  }
  return true;
};
```

Fig. 2. An event handler for the checkbox widget

Such a UI widget has a set of constraints on its accessibility states and state changes called an accessibility rule. A constraint on the accessibility states determines how to specify the accessibility state of a UI widget. For example, a checkbox widget must have a checked attribute and its value must be true or false. In addition, from those valid states we can define a constraint on state changes. For example, if there is a click event and the current value of a checked attribute is true, it should be changed to false. Otherwise it should be changed to true. We call such a set of constraints an accessibility rule. In the next section, we formalize the accessibility rule.

2.2 Formal Description for Accessible UI Widgets

We define an accessibility rule for each UI widget using a state transition diagram and a schema language playing a similar role with DTD based on regular expression type [3, 4] to describe the behavior of the UI widget and constraints on accessibility state respectively. In addition, the schema language is equivalent to the regular tree grammar [2].

For example, the following schema represents a checked checkbox where each attribute is considered as a label with prefixed @ and String represents an arbitrary string.

```
Checkbox → span[Id, Class, Checked]

Id       → @id[String]

Class    → @class[String]

Checked  → @checked[true]
```

This schema is corresponding to the following DTD definition.

```
<!ELEMENT  span  EMPTY!>
<!ATTLIST  span  id        #PCDATA!>
<!ATTLIST  span  class    #PCDATA!>
<!ATTLIST  span  checked  true!>
```

The syntax of the schema language is defined as follows.

```
Schema  := Rule, Rule, …
Rule     := Var → Tree
Tree     := Label [ Tree ]
          | Tree,Tree
          | Tree|Tree
          | Var
          | ()
          | {}
```

A schema consists of a set of rules. Each rule constrains the structure of a tree with a pattern Tree and is referred to by Var so that the pattern can be used in other rules. Such a pattern can represent a tree labeled by Label, a sequence of trees separated by "," or a choice of trees by "|", where the label represents HTML elements or attributes. In addition, "()" represents an empty sequence and "{}" means that there is no choice. An arbitrary number of trees is represented by "*", and it is defined as follows.

```
Var → Tree*  =  Var → Tree Var | ()
```

In our accessibility rules, the schema of the checked checkbox is generally defined by the following schema where Any represents an arbitrary element and A represents all the attributes except checked.

```
Checkbox → Any[A*, Checked, A*]

Checked  → @checked[true]

A        → (@id[String]|@class[String]|…)
```

If we define a schema for an unchecked checkbox, we use false instead of true in the rule for Checked.

The behavior of a UI widget is defined by a state transition diagram that consists of states and transitions. A transition has an event and a guard condition to determine when an accessibility state is changed. With a state transition diagram and schemata, we define the accessibility rules of UI widgets in which each state of a state transition diagram is associated with a schema representing an accessibility state.

Figure 3 shows an example of an accessibility rule for the checkbox. In the accessibility rule, CheckedSchema is a schema of the checked checkbox and it is associated with the checked state. UncheckedSchema is a schema of the unchecked checkbox and it is associated with the unchecked state. Each transition says that a checkbox changes its state when a user clicks the widget using a left button which is represented by 0 in JavaScript.

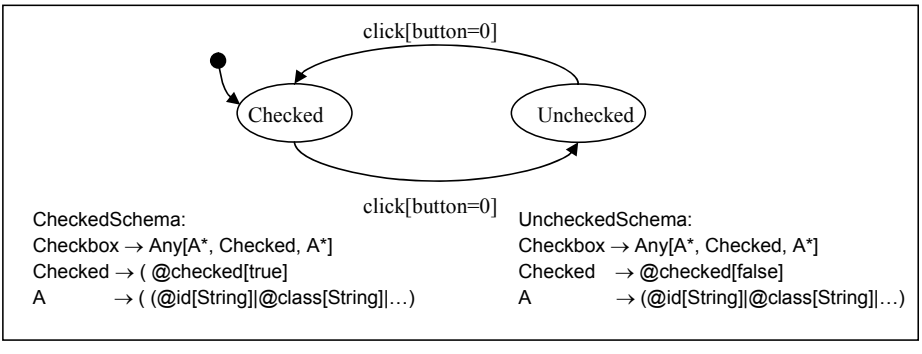


Fig. 3. State transition diagram representing behavior of the checkbox

3 Checking DHTML Accessibility

In this section, we describe how to check the accessibility of a UI widget against a corresponding accessibility rule. The main idea for this check is to verify that the UI widget is always accessible. More properly, it satisfies the corresponding schema as long as it behaves according to the corresponding state transition diagram. However the DOM instance representing a UI widget can be updated without limit and we cannot check all of the DOM instances. Therefore we focus on repeated updates and omit the checks against the repeated updates. Consider the example of the checkbox. The state of the checkbox is changed to checked and unchecked alternately as shown in Figure 4. Thus if we verify that (1) any DOM instance representing a checked checkbox is transformed to one representing a unchecked checkbox and (2) any DOM instance representing a unchecked checkbox is transformed to one representing a checked checkbox, we can omit the verifications of the other same transitions. This leads to the conclusion that we can divide the verification against the state transition diagram into one for each state transition, and check that each DOM instance satisfying a schema associated with a pre-state is transformed to one satisfying a

schema associated with a post-state. If all of the verifications of state transitions are successful, then the target UI widget satisfies the accessibility rule.

In order to describe the verification more precisely, we first briefly explain the static program analysis techniques used by the verification, and then describe how to use those techniques for the verification.

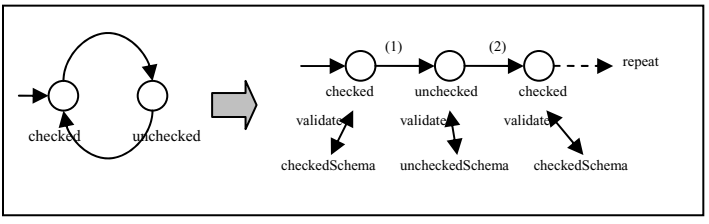


Fig. 4. Verifying an UI widget against an accessibility rule

3.1 Static Program Analyses

In the verification, we use the two complementary static program analysis techniques called dynamic slicing [8] and DOM analysis [7].

Dynamic Slicing

The dynamic slicing is a technique for extracting statements affecting the values of variables for particular input data. A collection of extracted statements is called a slice. As a simple example, consider a program shown in Figure 5 with input data event.type, event.button and event.target such that event.type is “click”, event.button is 0 and event.target is a DOM element that has a checked attribute with the value “true”. The deleted (struck-out) statements do not affect the DOM elements “checkbox” and we can extract the other statements as a slice.

```
if (event.type == "click" && event.button == 0) {
  var checkbox = event.target;
  if (checkbox.getAttribute("checked") == "true") {
     checkbox.setAttribute("checked", "false");
  } else {
     checkbox.setAttribute("checked","true");
  }
 return false;
}
return true;
```

Fig. 5. A slice of a JavaScript program

DOM Analysis

DOM analysis is a technique for inferring DOM instances as updated by a program. Such inferred DOM instances can be represented by a schema. Thus we can check that the inferred DOM instances are accessible by checking that they are included in the schema of an accessibility rule. For example, if we have a set of DOM instances

represented by the CheckedSchema schema of Figure 3 and infer the DOM instances as updated by the program of Figure 5, we can obtain a set of DOM instances represented by the following schema that represents a checked checkbox and an unchecked checkbox.

```
Checkbox → Any[A*, Checked, A*]

Checked  → @checked[true|false]

A        → (@id[String]|@class[String]|…)
```

In addition, if we use the sliced program shown in Figure 5, we can obtain a set of DOM instances represented by the following schema that represents only an unchecked checkbox.

```
Checkbox → Any[A*, Checked, A*]

Checked  → @checked[false]

A        → (@id[String]|@class[String]|…)
```

Those schemata can be compared with other schema such as CheckedSchema and UncheckedSchema of Figure 5 using the same algorithm as for regular expression types.

3.2 Verifying the Accessibility of UI Widgets

In order to verify the accessibility of a UI widget on a dynamic webpage against a corresponding accessibility rule, we first extract a DOM element providing a view of each UI widget and all event handlers that update the DOM instance (as described in Step 1), and then we check that the UI widget always satisfies a corresponding schema (in Step 2 and 3).

Step 1: Extract an UI widget and its event handlers
Accessibility roles are the key metadata for extracting UI widgets. We consider a DOM element having an accessibility role as a UI widget since assistive tools recognize classes of UI widgets based on the accessibility roles. In addition, we identify event handlers based on attributes such as onclick and onkeydown. For example, Figure 1 and Figure 2 are an extracted UI widget and the event handlers respectively, since the span element has the checkbox role and the onclick and onkeydown attributes are specified in the JavaScript program of Figure 2.

Step 2: Check that the initial DOM instance satisfies a schema
The main idea for checking accessibility of a DOM instance that could be updated without limit is to divide the verification against its corresponding accessibility rule into verifications against each state transition. This inductive idea is valid if the initial DOM instance is accessible. Thus we have to verify that the initial DOM instance extracted in Step 1 is accessible by comparing the DOM instance with the schema of the initial state. For the checkbox widget, we check if the DOM instance described in Figure 1 satisfies the CheckedSchema schema described in Figure 3.

Step 3: Check that every update is legal according to a corresponding state transition
We divide verification against an accessibility rule into verification against each state transition as described above. In the verification against a state transition, we verify

that an accessible DOM instance is transformed into another accessible one with a corresponding event handler.

For example, when we verify that the checkbox described in Figures 1 and 2 is updated according to the state transition from the checked state to the unchecked state shown in Figure 3, we can extract the same slice as in Figure 5. This is because the state transition has a click event and the guard condition button=0 and thus we can use "click" and 0 for event.target and event.button respectively. Next we can infer a set of DOM instances from the CheckedSchema schema. The set of the inferred DOM instances are represented by the following schema.

```
Checkbox → Any[A*, Checked, A*]

Checked  → @checked[false]

A        → (@id[String]|@class[String]|…)
```

Finally we check that the set of the inferred DOM instances is included in the UncheckedSchema schema.

For this verification, we perform the following three steps:

1. Extract a slice corresponding to the target state transition from the event handler. In order to extract a slice corresponding to a target state transition, we use an event, a guard condition and a schema associated with a pre-state to identify the input data.
2. Infer a set of updated DOM instances from a set of DOM instances that satisfies the schema associated with the pre-state of the state transition by applying DOM analysis to the extracted slice. Note that the set of the inferred DOM instances contains all of DOM instances arising at run-time and it is represented in the schema language.
3. Check that the set of the updated DOM instances is included in the corresponding schema associated with the post-state of the state transition.

4 Experiments

We developed a prototype of a DHTML accessibility checker based on the proposed technique. The prototype has prior knowledge of the accessibility rules corresponding to the UI widgets. Thus developers need not prepare any accessibility rules by themselves.

We used the prototype for checking the accessibility of a checkbox in accessible DHTML samples provided by the Mozilla.org project against three types of the accessibility rules shown in Figures 3, 6 and 7, testing on a laptop PC with a Pentium-M 2.13GHz CPU and 1.5GB of memory. The time for the slicing steps (Step 3-1), the inference steps (Step 3-2), the comparison steps (Step 3-3) and the total time are shown in Table 1.

As shown in the results, most of time was taken by the comparison phases, since we employed only the basic algorithm for comparing regular expression types though there are some high-level techniques to improve performance. In addition, theoretical total time for inference phases and comparison phases proportionally increases according to the number of state transitions. This prospect is shown in the results of Figure 3 and 6. However, the inference time and the comparison time for Figure 7 is worse than Figure 3 since an extracted slice and a set of inferred DOM instances for Figure 7 become larger or more complex than Figure 3.

Table 1. Execution time of the verification

accessibility rule	slicing time (ms)	inference time (ms)	comparison time (ms)	total time (ms)
Figure 3	234	516	3860	7391
Figure 6	251	953	6186	10140
Figure 7	203	750	24625	28266

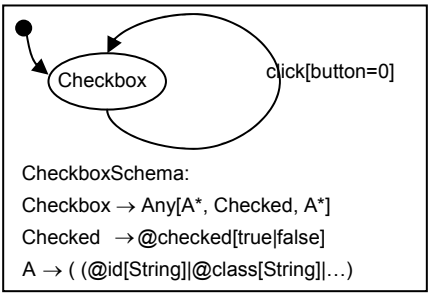


Fig. 6. An accessibility rule with one state

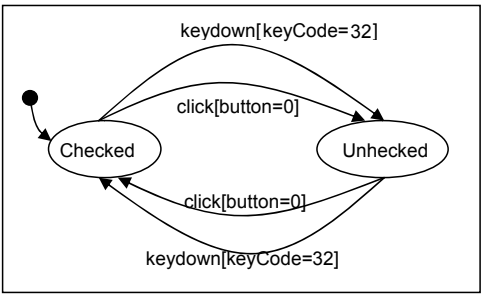


Fig. 7. An accessibility rule with the keydown event

5 Related Work

There are a lot of automated accessibility check tools [16] such as WebXACT [17] or WebKing [9] and related research activity. As for WebKing, it provides functionality for checking dynamically constructed webpages. For this check, it requires scripts in which the tester describe how a user reacts to the webpage, such as by filling out an input form with text, and then it checks the pre-defined rules, including the accessibility by executing the applications. On the other hand, our proposed technique does not need to execute so it is applicable to accessibility checking during the development phase.

From the viewpoint of the flexibility of tools, a concept of separation between checking rules (i.e.: standards and guidelines) and rule engines (i.e.: system for evaluating the checking rules) is important. Some tools such as Lift [12] allow us to customize a guideline according to specific requests such as a corporate guideline by selecting pre-defined rules. In order to make a customization more flexible, a language for defining guidelines and a framework for evaluating custom guidelines was proposed [1, 6]. We can describe static features of guidelines using the guideline definition language, but we can not describe dynamic features such as the behavior of the UI widget.

In addition, Sun et al. [11] and Pontelli et al. [10] focus on semantic features of webpages for improving navigability. Sun uses automata for modeling online web transactions as we use state transition diagrams for modeling the behavior of the UI

widgets. Pontelli provides a language for providing semantic description of tables found in web pages and for navigation strategies.

6 Conclusion

In this paper, we proposed a fundamental technique for verifying the accessibility of UI widgets in a DHTML application during the development phase. In this verification, we check that an UI widget always has proper accessibility state according to a corresponding accessibility rule. In addition, we use power of static program analysis techniques so that we can perform the accessibility checking without execution. In the future, we will improve our prototype so as to apply it to a set of UI widgets implemented as a part of JavaScript libraries such as DOJO.

References

1. Beirekdar, A., Keita, M., Noirhomme, M., Randolet, F., Vanderdonckt, J., Mariage, C.: Flexible Reporting for Automated Usability and Accessibility Evaluation of Web Sites. In: Proc. of Human-Computer Interaction (2005)
2. Comon, H., Dauchet, M., Gilleron, R., Jacquemard, F., Lugiez, D., Tison, S., Tommasi, M.: Tree Automata Techniques and Applications http://www.grappa.univ-lille3.fr/tata
3. Hosoya, H., Vouillon, J., Pierce, B.: Regular expression types for XML. In: Proc. of the International Conference on Functional Programming, pp. 11–22 (2000)
4. Hosoya, H., Pierce, B.: Regular expression pattern matching for XML. In: Proc. of Principles of Programming Languages, pp. 67–80 (2001)
5. Mozilla: Accessible Rich Internet Application http://developer.mozilla.org/en/docs/Accessible_DHTML
6. Vanderdonckt, J., Beirekdar, A.: Automated Web Evaluation by Guideline Review. Journal of Web. Engineering 4(2), 102–117 (2005)
7. Tateishi, T., Miyashita, H., Saito, S., Ono K.: Automated Verification Tool for DHTML. In: Proc. of Automated Software Engineering (2006)
8. Tip, F.: A survey of program slicing techniques, Journal of Programming Languages, pp. 121–181 (1995)
9. Parasoft: WebKing, http://www.parasoft.com/jsp/products/home.jsp?product=WebKing
10. Pontelli, E., Xiong, E., Gupta, G., Karshmer, A.I.: A Domain Specific Language Framework for Non-Visual Browsing of Complex HTML Structures. In: Proc. of the International Conference on Assistive technologies (2000)
11. Sun, Z., Mahmud, J., Mukherjee, S., Ramakrishnan, I.V.: Model-directed web transactions under constrained modalities. In: Proc. of the International Conference on World Wide Web, pp. 447–456 (2006)
12. UsableNet Inc.: Lift for Dreamweaver http://www.usablenet.com/products_services/lift_dw/lift_dw.html
13. W3C: Dynamic Accessible Web Contents Roadmap http://www.w3.org/WAI/PF/roadmap/
14. W3C: Roles for Accessible Rich Internet Applications http://www.w3.org/TR/aria-role/
15. W3C: States and Properties Module for Accessible Rich Internet Applications http://www.w3.org/TR/aria-state/
16. W3C: Web Accessibility Evaluation Tools http://www.w3.org/WAI/ER/tools/
17. Watchfire Corporation: WebXACT http://webxact.watchfire.com/

A Unified Web Evaluation Methodology Using WCAG

Eric Velleman[1], Christophe Strobbe[2], Johannes Koch[3], Carlos A. Velasco[3], and Mikael Snaprud[4]

[1] Accessibility Foundation, Oudenoord 325, 3513EP Utrecht, Netherlands
e.velleman@bartimeus.nl
[2] Katholieke Universiteit Leuven, Kasteelpark Arenberg 10, B-3001 Heverlee-Leuven, Belgium
{Christophe.Strobbe,Jan.Engelen}@esat.kuleuven.be
[3] Fraunhofer-Institut für Angewandte Informationstechnik (FIT), Schloss Birlinghoven, D-53757 Sankt Augustin, Germany
{Johannes.Koch,Carlos.Velasco}@fit.fraunhofer.de
[4] Agder University College, Serviceboks 422, N-4604 Kristiansand, Norway
mikael.snaprud@hia.no

Abstract. Checking of web accessibility can be carried out in several ways along the same international standards and depending on the scale, the quality, availability of tools and the interpretation method that is followed. The Unified Web Evaluation Methodology is developed by European expert organizations and offers test descriptions to evaluate WCAG 1.0 conformance covering level AA, a clear sampling scheme, improved aggregations supporting confidence levels, Score cards and other instruments to help communicate the results of evaluations clearer and on more levels. The aim is to establish the UWEM as the basis for web accessibility evaluation, policy support and possible certification in Europe.

Keywords: web accessibility, conformance, certification, unified web evaluation methodology, uwem.

1 Introduction

Accessibility checks of webresources can be carried out in different ways even if the checks are based on the same internationally accepted guidelines. The Unified Web Evaluation Methodology (UWEM) is developed to ensure that large scale monitoring and local evaluation are compatible and coherent among themselves and with the Web Content Accessibility Guidelines from W3C/WAI [1]. The methodology is based on the Web Content Accessibility Guidelines of the World Wide Web Consortium (W3C) Web Accessibility Initiative (WAI), and will be synchronised with the foreseen migration from WCAG 1.0 to WCAG 2.0 [2]. Thus, the UWEM will sharpen the support for evaluating, certifying, and benchmarking web content in conformance with WCAG.

The Unified Web Evaluation Methodology offers test descriptions to evaluate WCAG 1.0 conformance covering level AA, a clear sampling scheme, improved

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 177–184, 2007.
© Springer-Verlag Berlin Heidelberg 2007

aggregations supporting confidence levels, Score cards and other instruments to help communicate the results of evaluations clearer and on more levels. The aim is to establish the UWEM as the basis for web accessibility evaluation, policy support and possible certification in Europe. In a recent communication on eAccessibility, the European Commission announces they will foster the use of three approaches not yet widely used in Europe: (i) accessibility requirements in public procurement, (ii) accessibility certification, and (iii) better use of existing legislation [3]. The UWEM aims to support these approaches.

2 The WAB Cluster

The UWEM has been produced from the Web Accessibility Benchmarking Cluster (WAB Cluster). The work of the Web Evaluation Benchmarking Cluster involves three European projects with a total of more than 21 partners from eight countries. They work together on harmonisation and evaluation to provide a European base for evaluation of Websites for accessibility. The projects participating in the WAB cluster are funded by the European Union in the second FP6 IST call (2003) of the eInclusion Strategic Objective. More information about the projects can be found on the project web sites:

http://www.wabcluster.org
http://bentoweb.org
http://www.eiao.net
http://www.support-eam.org

The Ben To Web project supports the European public and private sector by providing new software modules and methodologies for Web Accessibility, which are not analysed by existing tools due to their inherent complexity. BenToWeb also produces suites of test files for WCAG 2.0, which can be used for benchmarking accessibility evaluation and repair tools. More information: http://www.bentoweb.org

The EIAO project establishes the measurement machinery for a prototype EU Internet Accessibility Observatory. Frequently updated measurements will be available online from a data warehouse to enable large scale benchmarking, and provide a basis for policy-making, research and actions to improve the accessibility to Internet content. More information: http://www.eiao.net.

The Support EAM project has produced a CEN Workshop Agreement [4] on certification, many evaluation materials, tutorials and a curriculum. It has put down the basis for the use of the UWEM as a framework for a unified inspection process as part of a European certification scheme for Web accessibility based on interaction with users [5]. More information: http://www.support-eam.org

3 UWEM, Aims and Requirements

The Unified Web Evaluation Methodology provides an evaluation procedure offering a system of principles and practices for expert and automatic evaluation of Web accessibility for humans and machine interfaces. The Methodology is designed to be

conformant with WCAG 1.0 priority 1 and 2 checkpoints with regard to technical criteria.

The UWEM aims to increase the value of evaluations by basing them on a shared interpretation of WCAG 1.0 and a set of tests that are sufficiently robust to give stakeholders confidence in results. Web content producers may also wish to evaluate their own content and the UWEM aims to also be suitable to these users.

The methodology is designed to meet the following requirements:

- Technical conformance to existing Web Accessibility Initiative (WAI) Recommendations and Techniques documents;
- Tool and browser independence: questions and tests are given in a 'pure' form, making them as tool independent as possible;
- Unique interpretation: questions shall have only one way of being interpreted;
- Replicability: different Web accessibility evaluators who perform the same tests on the same site should get the same results within a given tolerance;
- Translatability: the methodology will address localization issues;
- Compliance with Regulation (EC) No 808/2004 of the European Parliament and of the Council of 21 April 2004 concerning Community statistics on the information society.

In the methodology information has also been included about:

- Scope and sampling;
- Reporting, interpretation and integration/aggregation of results.

The UWEM primarily covers methods to evaluate documents based on the following technologies:

- HTML 4.01
- XHTML 1.0 and 1.1
- CSS 2.x
- Other embedded objects in (X)HTML resources

4 Target Audience of the UWEM

The target audiences for the UWEM include, but are not limited to:

- Web accessibility benchmarking projects
- Possible Certification Authorities
- Web content producers wishing to evaluate their content
- Developers of Evaluation and Repair Tools
- Policy makers and Web site owners
- Other organizations evaluating Web sites

The European Commission, national governments and other organizations who wish to benchmark Web accessibility will be able to use the UWEM to carry out the

evaluations and compare their results in a meaningful way. For this reason, a score-card solution is presented.

The UWEM is an evaluation methodology and is not intended to provide information for Web content producers wishing to produce content conformant with WCAG 1.0. That information is provided in the WCAG 1.0 Techniques Documents [6] that are available through the W3C/WAI web site.

5 Automatic, Expert and User Testing

Web accessibility may be tested through automatic, expert and user testing. Some companies evaluating websites for accessibility on a commercial basis only use manual evaluation while others use only automatic testing or a combination of both [5]. Currently expert testing and automatic testing are supported in the UWEM. A section including testing protocols for specific web evaluation tests with people with disabilities has been delivered in an earlier stage. The testing protocols are used in the BenToWeb project.

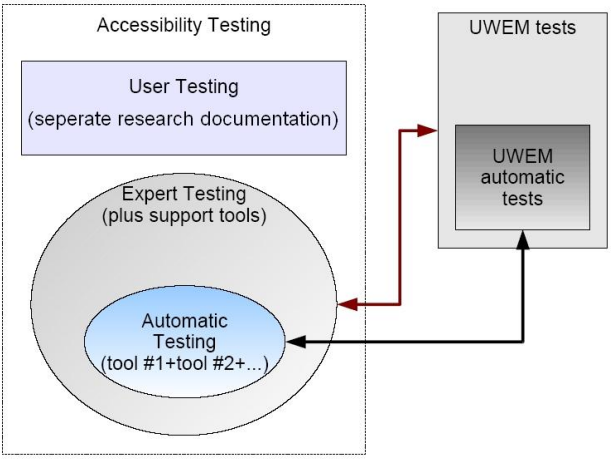


Fig. 1. The different evaluation methods and the coverage of the UWEM

Figure 1 describes three different evaluation methods. Currently two (automatic and expert) are covered in the UWEM. Automatic testing of all checkpoints for accessibility is not possible. If a checkpoint is fully automatable, this is indicated. Figure 1 shows that automatic tool #1 and #2 can only identify a subset of all possible accessibility barriers. Tools better support testing of very large numbers of Web resources within one or multiple Web sites. Some tools can also act as support systems in an expert evaluation process. The tools provide reliable results for a subset of tests and can not only speed up the process by performing some tasks automatically, but also, by providing hints about barrier locations, indicate areas the expert evaluators should focus on.

User testing is able to identify barriers that are not caught by other testing means, and is also capable of estimating the accessibility for tested scenarios. However, user testing is quite specialised, thus it is not generally suitable for conformance testing. The best approach to ensure both accessibility and the UWEM conformance is to use a combined approach encompassing all evaluation methods: automatic, expert evaluation and user testing of the Web site. Involvement of users in the evaluation of web content is described in the W3C/WAI evaluation suite (available on the WAI website).

6 Claiming Conformance

The purpose of the UWEM is to guarantee replicability of results. The unambiguous expression of the tested resources is of key importance for the aggregation and comparison of results. To claim conformance with the UWEM, it is minimally required that the resources sample and the scope for the evaluation are defined and that all resources in the sample should pass all applicable tests to the corresponding conformance level. The size and selection of samples is also described in the UWEM.

The conformance levels to the UWEM replicate those of WCAG10 [1]. The claims of accessibility conformance for manual evaluation according to the UWEM methodology must use the following form:

1. The UWEM version and its URI identifier, i.e.;
2. The URI to a document detailing the scope and the sample to which the claim refers;
3. The level of conformance.

Evaluation tools can also claim conformance to the UWEM 1.0. In that way, experts evaluating web sites according to the UWEM 1.0 will be able to rely on the results of the tool for the fully automatable tests of the methodology. To claim conformance to the UWEM 1.0, the tool MUST implement all fully automatable tests to the corresponding conformance level.

7 Tests and Structure of Tests for Conformance Evaluation

The UWEM offers tests for automatic and expert testing of Priority 1 and Priority 2 checkpoints of WCAG 1.0. It does not repeat information available in W3C documents. Instead, it provides pointers to the relevant W3C information if available and extends only information when necessary for the defined tests. Web content passes a checkpoint if it fails none of the applicable tests for that checkpoint. Web content fails a checkpoint if it fails any of the applicable tests for that checkpoint.

The structure of the tests is harmonized throughout the document and based on extensive evaluation experience by the partners in the projects:

1. Guideline: Quotation of the corresponding WCAG 1.0 guideline with pointers to additional clarifications on W3C website if available.

[1] http://www.w3.org/TR/WCAG10/#priorities

2. Checkpoint: Quotation of the corresponding WCAG 1.0 checkpoint. Pointers to additional clarifications on W3C website if available. For each checkpoint, a set of one or more tests is defined. If no automatic tests for a certain technology are defined, this means that there are no applicable tests for automated testing.
3. Technology specific tests: There are tests for (X)HTML resources and other technologies. Each test consists of:

- **Title and ID:** short descriptive title (informative) and unique identifier (normative);
- **Applicability criteria:** elements, attributes and combinations thereof used to determine the applicability of the test. Whenever possible, the criteria are presented as XPath expressions, otherwise a prose description is given;
- **Test procedure:** description in a tool-independent manner of the test procedure. The procedure may consist of multiple steps and is written so as to enable possible machine-testing;
- **Expected results:** statement defining the fail or pass conditions with regard to one or more steps in the test procedure. The elements or content specified in the accessibility criteria pass the test if the result is not FAIL;
- **Fully automatable:** statement whether the test procedure can be fully automated (yes/no).

8 Aggregation and Reporting of Test Results

The UWEM provides models for aggregation and reporting of test results including a User Centric Accessibility Barrier model, EARL schema, an Evaluation Report Template and a Scorecard that gives easy understanding and clear interpretation to the end-result. Following the European Commission regulation 808/2004 concerning community statistics for the information society, the UWEM provides a model for calculating the accessibility barrier probability for single Web pages and Web sites. A single value can then describe the probability of creating a barrier by violating any of the evaluation tests.

Scorecards	RGB	Barrier probalility Fs	Level of accessibility	Interpretation
A	204/236/255	Fs = 0.00%	Accessibility is fully supported	Full accessibility
B	153/204/255	0% < Fs <= 25%	Within levels of confidence no or very few accessibility barriers have been identified	Very good accessibility
C	0/153/204	25% < Fs <= 50%	A few accessibility barriers have been identified	Medium accessibility
D	51/102/153	0% < Fs <= 75%	Several accessibility barriers have been identified	Poor accessibility
E	0/51/102	75% < Fs <= 100%	Web Site is inaccessible	Very poor accessibility
n/a	255/255/255	Not measured	Not tested or not applicable	n/a

Fig. 2. UWEM barrier scorecard scores represented by letters and colours to indicate barrier probability. Note that the levels chosen are preliminary values that will be evaluated when real data measurements from the EIAO project are available.

The scorecards are based on this value that describes the probability of creating a barrier. The scorecards can be used to monitor the status and progress of single Web sites and groups of Web sites over time.

The scorecard gives a helicopter view of the status and progress, and allows to monitor a set of indicators. Figure 2 shows the colour and letter codes used for the UWEM barrier status score. The coloured codes were based on the shades of blue as this is a set of colours that ensures a clear distinction regardless of colour-vision impairment and on the black-and-white printout.

The Manual Evaluation Report Template included in the UWEM closely follows the W3C evaluation suite template and offers evaluators a clear template to report accessibility.

9 Conclusion and Future Work

The Unified Web Evaluation Methodology aims to increase the value of evaluations by a shared interpretation of WCAG 1.0 and a set of tests that are sufficiently robust to give stakeholders confidence in the results. It offers detailed test descriptions covering WCAG1.0 level AA, a clear sampling scheme, improved aggregations supporting confidence levels, Score cards and other instruments to help communicate the results of evaluations clearer and on more levels.

There are parts of the methodology where more information and research is needed. This requires more evaluation results than currently available. More detailed recommendations will be added in the next version of the UWEM and will be largely based on the results from experiments within the EIAO and BenToWeb projects. Also more public and user comments are necessary. A possible channel for that is provided through the forum on www.wabcluster.org/forum/. Based on these evaluations and experiments, a next version will be made available that includes support for WCAG2.0. This next version will also include more information on user testing, aggregation, reporting and sampling information and offer the possibility to choose the WCAG version.

The current version of the UWEM has been used extensively by different European organization for commercial testing of websites. Already a French and a Spanish translation exist (www.wabcluster.org).

Acknowledgements. This paper is based on work done on the Unified Web Evaluation Methodology by all partners in the Web Accessibility Benchmarking Cluster. The persons directly involved in the UWEM are: Eric Velleman (Clusterleader); Carlos A Velasco; Mikael Snaprud; Dominique Burger; Jesse Tan; Colin Meerveld; Barry McMullin; Nils Ulltveit-Moe; Johannes Koch; Annika Nietzio; Helen Petrie; Jenny Craven; Alan Chuter; Gerhard Weber; Christophe Strobbe; Sandor Herramhof; Evangelos Vlachogiannis; Pierre Guillou; Alistair Garrison; Olaf Perlick; Yehya Mohamad; Henrike Gappa; Gabriele Nordbrock; Dirk Stegemann; Finn Aslaksen; Torben Bach Pedersen; Patrizia Bertini; Terje Gjøsæter; Morten Goodwin Olsen; Ole-Christoffer Granmo; Andreas Prinz; Agata Sawicka; Edith Helene Unander.

This project has been undertaken in the framework of the project BenToWeb IST-2-004275-STP funded by the IST Programme of the European Commission and involves the partners working in the projects EIAO and Support EAM.

References

1. W3C Recommendation. Web Content Accessibility Guidelines 1.0 (WCAG 1.0). In: Chisholm, W., Vanderheiden, G., Jacobs, I(eds.) World Wide Web Consortium (MIT, INRIA, Keio) (May 5, 1999) available at http://www.w3.org/TR/WCAG10/
2. W3C Working Draft. Web Content Accessibility Guidelines 2.0 (WCAG 2.0). In: Caldwell, B., Chisholm, W., Vanderheiden, G., White, J. (eds.) World Wide Web Consortium (MIT, ERCIM, Keio) (2006) available at http://www.w3.org/TR/WCAG20/
3. European Commission: Communication COM, 425 - Communication from the Commission to the Council, the European Parliament, the Economic and Social Committee and the Committee of Regions. Brussels (September 13, 2005) (2005) also available at: http://europa.eu.int/information_society/policy/accessibility/com_ea_2005/a_documents/cec_com_eacc_2005.html
4. CWA 15554, Specifications for a Web Accessibility Conformity Assessment Scheme and a Web Accessibility Quality Mark. CEN (2006), available at ftp://ftp.cenorm.be/PUBLIC/CWAs/e-Europe/WAC/CWA15554-00-2006-Jun.pdf
5. Support, EAM.: deliverable D3.2 Stakeholders' needs for a Quality Mark (April 8, 2005) also available at http://www.support-eam.org
6. W3C note. Techniques for Web Content Accessibility Guidelines 1.0. In: Chisholm, W., Vanderheiden, G., Jacobs, I. (eds.) W3C Note. World Wide Web Consortium (November 6, 2000) (1999) Available at: http://www.w3.org/TR/WAI-WEBCONTENT-TECHS/
7. Web Accessibility Benchmarking Cluster: D-WAB2 unified web evaluation methodology (UWEM 1.0) (2006) available from http://www.wabcluster.org/uwem/

Improving Accessibility Through the Visual Structure of Web Contents

Masahiro Watanabe, Daisuke Asai, and Yoko Asano

Cyber Solutions Labs., Nippon Telegraph and Telephone Corp.,
1-1 Hikari-no-oka, Yokosuka-shi, Kanagawa, Japan
{watanabe.masahiro,asai.daisuke,asano.yoko}@lab.ntt.co.jp

Abstract. Web accessibility has become one of the most popular research targets. Web accessibility means the ability to be accessed by all kinds of people or devices. One problem that some Web pages have rather poor information structures at the HTML source code level, even though the pages are well structured visually. In order to transfer the visual structure to the HTML level structure, we need to identify the graphic design features that influence human understanding. We examined whether several people interpreted a displayed structure in the same way or not. Four subjects participated in trials; they attempted to identify the headers, major items forming the structure of the top page, of the web sites of 7 local governments in Japan. 80 % of the headers were selected by all subjects so most subjects shared the same understanding of what constituted a header.

Keywords: Web, accessibility, information structure, visual structure.

1 Introduction

As Internet access has become more popular, Web sites have become one of the most important methods of distributing information. Therefore, Web contents should be accessed by all kinds of people and devices. A good recommendation for making Web content accessible is to follow effective design guidelines. Accessibility guidelines for people with disabilities and the elderly, Web Content Accessibility Guidelines 1.0, WCAG 1.0 were established by World Wide Web Consortium in 1999 [1]. Since Japanese society is aging rapidly, the concept of Web accessibility has become more and more important. The Japanese Industrial Standard also established guidelines in 2004 [2]. It is now a legal requirement, especially for the Web sites of public bodies such as local governments. We have made our guidelines and encouraged their use throughout our group companies [3, 4].

Owing to progress in assistive technology, Web pages can now be presented not only visually (conventional manner) but also aurally [5]. A common practice is to create different material for visual and aural presentation. This is inefficient as it is difficult to update both contents at the same time. The ideal web site has only one information source but various presentation methods to suit different users and their situations as shown in Fig. 1. This is the universal design concept in Information

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 185–192, 2007.
© Springer-Verlag Berlin Heidelberg 2007

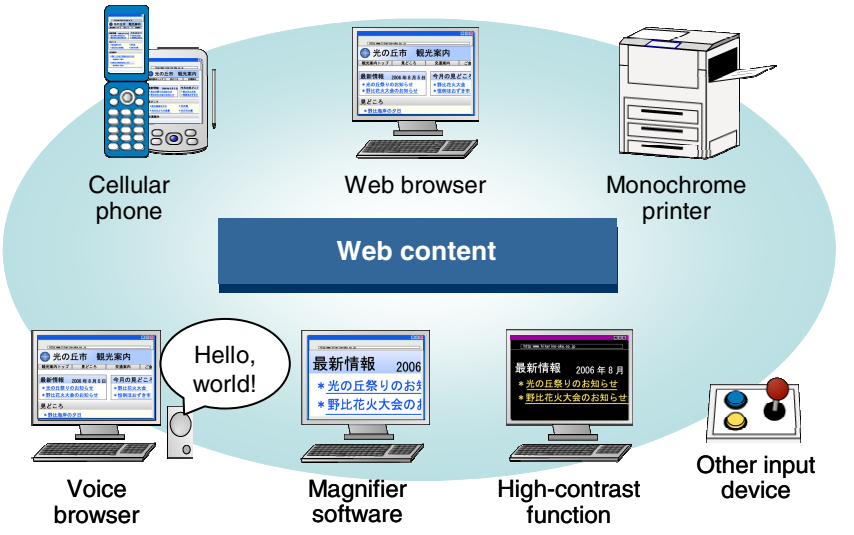


Fig. 1. Concept of universal design in Web content. The ideal web site has only one information source but various presentation methods to suit different users and their situations.

Technology. One problem that some Web pages have is to transfer efficiently the content according to the kind of user agent. User agent means software by which we can browse Web content such as the common Web browser, voice browser, and so on.

In order to transfer Web content according to user agent, the information structure of the Web content itself is especially important, not only for Web sites but also Web pages. Web content is written in (X)HTML which defines the content's structure. If content satisfies HTML grammar and is well structured, it can be well used by a wide variety of user agents including voice browsers. Information structure is the key to Web accessibility. In this paper, we focus on the visual information structure in each Web page. We present a method that assists in the creation of accessible Web pages based on visual structure.

2 Visual Structure and HTML Source Code Structure

If information structure is correctly described in HTML source code, it also present well visually. For example, a word tagged H, header element in HTML source code, is presented bigger than the other words and inserts a line feed. Unfortunately, these features are misused in some Web pages, for example, they use the H element just to increase word size. This is because why some existing guidelines such as WCAG 1.0 suggest using the style-sheet technique to control the appearance. In an ironic twist of fate, the style-sheet technique can yield Web contents that have no information structure at the HTML source code level. We need a well-defined structure at the HTML source code level if assistive technologies such as voice browsers are to present the content efficiently.

In contrast, the visual designs used to present Web content have really become sophisticated recently, because Web sites have become so popular. One problem is that some Web pages have rather poor information structures at the HTML source code level, even though they are well structured visually. This separation needs to be examined and countered.

Our basic idea is to use the visual structure to elucidate the HTML source structure. To do this we need to identify the graphic design features that influence human understanding. Examples include font size, color, and background color. A good visual structure allows all readers to recognize the same part as a header.

To extract information structure from a Web page makes it possible to create new technique that permits users to browse the page easily and efficiently. Extracting the information structure of a Web page has become a popular research target. Some studies proposed extraction methods based on Document Object Model, DOM [6, 7]. Unfortunately, many existing Web sites have poor structure and it is not until the information structure in a Web page is described correctly that the right DOM structure is used. Only if Web sites that have correct information structures become popular will these studies have practical meaning. Some techniques for assisting the design of Web pages that have correct information structures and that satisfy HTML grammar should be developed. Some studies deal with segmentation methods based on pixel position of the words on the page [8]. Other studies deal with the user's expectation of Web elements' locations on a Web page [9]. These studies will contribute to the progress of accessibility techniques.

In this paper, we focus on the H element for discerning the structure of Web pages. The H element in HTML source code is one of the most important elements to show the structure of a Web page. Some guidelines such as WCAG 1.0 note the importance of designing Web page with H elements.

3 Survey of Information Structure in HTML Source Code

We examined the HTML source files of the Web sites of 47 local governments in Japan in order to investigate whether H elements are popular or not. We looked for headers in the HTML source code of these Web pages.

The result shows only 27 sites (57 %) used one or more H elements in the top pages as shown in Table 1. However, only 7 of these sites had correct information structures at the HTML source code level.

The headers illustrated in Fig. 2 were found to be quite common. These results suggest that there are many Web sites that are well structured visually but not at the HTML source code level. These results also show that almost the half sites examined paid little attention to Web accessibility.

Table 1. Existance of H elements in Japanese local governments' Web sites

Sites' types	Number of headers	Percentage
At least one H element used	27	57
No H elements used	20	43
Total	47	100

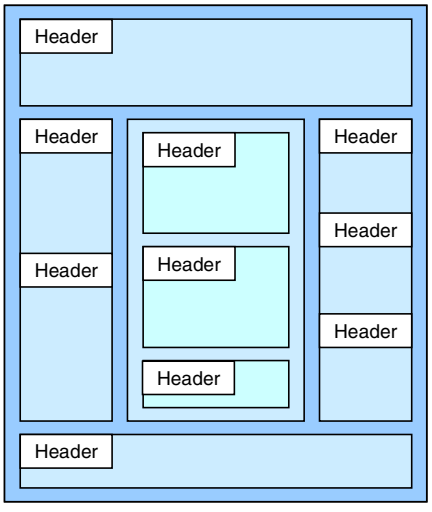


Fig. 2. Visual information structure of a Web page

4 Experiment on Visual Information Structure

4.1 Outline of the Experiment

We examined the agreement between the judgments of several people as to the displayed structure of a Web page. We then examined whether the displayed structure was accepted by several people in the same way or not.

We investigated the 7 top pages of the Web sites of local governments in Japan (Chiba, Gunma, Kanagawa, Saitama, Tochigi, and Tokyo). Four subjects (two females; ages 24 - 40) who use Web browsers in their jobs participated in this experiment; they attempted to identify the headers, major items forming page structure, of the top page, of 7 local governments' sites in Japan as shown in the next section.

4.2 Procedure

The main purpose of the header is to clearly indicate the meaning of each component in a Web page. The subjects separated each Web page into several components and for each component selected the expression that described the component. We defined expression as the "recognized header". For each component, the subjects were told to identify sub-components and the recognized header of each sub-component. They repeated this until no further recognized headers could be found.

First of all, we printed the Web pages adjusted to A4 sheets. The subjects used the procedure shown below to pick up the recognized headers on the printed Web top page. In this experiment, the subject was told to separate the components according to graphic design, not meaning.

1. Separate a Web page into some sub-components, which include as few similar items in each component as possible.
2. Mark only one expression as most representative of the component.
3. Select items from questionnaires as to why the subject marked the expression and put them the order of agreement.
4. Separate each component except the marked expressions into some sub-components similar to procedure 1.
5. Quit this procedure if the subject feels there is no expression that well represents for the component after separating the component, or there is only one expression in a component.

We regarded the expression marked in procedure 2 as a recognized header. We prepared the choices in procedure 3 based on some Web design text books as shown below. If the subjects selected "Others", they were asked to detail their reasoning.

1. Size of characters: bigger than others
2. Color of characters: different from others
3. Font type: different from others
4. Bold characters
5. Icon in front of characters
6. Color of characters' background: different from others.
7. Frame around characters
8. Line separates the word
9. Wide space around the characters
10. Indent of the word
11. Unindent of the word
12. Lists below the word
13. Others

4.3 Results

We investigated the individual differences in selecting recognized headers among the subjects. Table 2 shows how many recognized headers were selected by subjects with the same understanding. As shown in Table 2, 194 expressions were selected by at least one subject. We considered the 194 expressions as recognized headers. All four subjects selected 155 expressions (80 %) as recognized headers so they shared the same understanding of what constituted a recognized header. This result implies that

Table 2. Individual differences among visually indicated headers

Number of subjects with same understanding	Number of headers	Percentage
4	155	80
3	6	3
2	8	4
1	25	13
Total	194	100

there was only a small difference among the subjects and the headers succeeded in being accepted as headers by all subjects.

In further analyses, we used the 155 recognized headers accepted by 4 subjects. We compared the 155 recognized headers to the headers identified by the H element in HTML source code in order to clarify the adequacy of their information structures. Only 4 Web pages among 7 pages used at least one H elements. There are 78 recognized headers in the 4 pages. Table 3 shows how many recognized headers were tagged in HTML source code. It shows that only 60 % of the recognized headers were supported by correct information structures at the HTML source code level. There are at least two reasons why an expression was accepted as a recognized header. An average of 3.3 features were selected as the reasons for recognized header idenitification.

Table 3. Proportion of adequatly tagged recognized headers

Recognized headers' types	Number of headers	Percentage
Tagged in HTML source	47	60
NOT tagged in HTML source	31	40
Total	78	100

We then asked the subjects to explain their selection by selecting from among 12 explanations. Of the 155 recognized headers selected by all subjects, 85 of them (55%) had at least one explanation that was selected by all subjects as shown in Table 4. Table 4 shows the proportion of the graphic features which were selected by the subjects. The proportion of the graphic features that were selected by all of the subjects was 55 % of the all features which were selected by at least one subject.

Table 4. Proportion of graphic features selected by the subjects

Number of subjects selecting the header for the same reason	Number of headers	Percentage
4	85	55
3	57	37
2	13	8
1	0	0
Total	155	100

These results imply that all subjects had almost the same approach to header recognition. Further research is needed to precisely identify exactly which graphic features are important. By the way, the order of subjective evaluation was not the same so we did not analyze the order precisely.

It is interesting to note that 85 recognized headers were selected by all subjects with only 5 features as shown in Table 5. Table 5 shows the graphic features selected by all subjects. Not that this table shows the number of recognized headers according to each feature. Two recognized headers were selected with 2 features each. This is why there are 87 recognized headers in Table 5.

This result shows that the color of character or background has a great impact. It is thought that site designers find it easy to change color as a feature. On the contrary, if the characters were bold, the subjects might select different feature such as size, bold, or font. The choice should not be ambiguous.

Table 5. Features for headers identification

Features	Number of headers	Percentage
Color of characters	36	41
Color of characters' background	30	34
Size of characters	13	15
Icon in front of characters	7	8
Unindent of the word	1	2
Total	87	100

5 Discussion

The result of the survey described in Section 3, shows that 43 % of Web pages do not have any H tags. It means that these Web pages have no information structures at the HTML source code level. This suggests that these sites will not be well handled by voice browsers. Although the Web sites of Japanese local governments follow the guidelines of Japan Industry Standards and have high levels of accessibility, almost half of them have no HTML level structures. This problem is likely to be wide spread. On the contrary, visual structures were recognized in all 47 Web pages examined. This implies that Web designers succeeded in making Web page whose visual structures were easily understood. If they paid more attention to recognized headers and adjusted the structure at the HTML level, the content would become accessible.

The results of experiment, see Section 4, show that there are only small individual differences in recognized headers selection. This suggests that visual design is really sophisticated and users can easily understand the content's structure at the visual level; recognized headers can be identified even if they are not tagged. This indicates the possibility of new header extracting techniques based on human understanding. This idea can also be used to enhance accessibility evaluation tools that are intended to help Web designers to make Web content accessible. If Web sites that have correct information structures become popular, several techniques based on DOM will become effective. To achieve this, we have to investigate the visual features which impact human understanding on Web content in detail. In this experiment, we focused on character features such as font size, color, and background color. We should investigate not only the visual feature of the header itself but also the layout of the header on the page. The difference between the features should be studied in detail.

6 Conclusion

We examined the HTML source files of the Web sites of 47 local governments in Japan and looked for the headers in the HTML source code in these Web pages. Our

results show that only 57 % of the sites used the H element in the top pages which indicates that many Web sites have poor information structures, even though they have good visual structure.

We also examined the agreement between the judgments of several people as to the displayed structure of 7 top pages of Web sites of local governments in Japan. The result showed that 80 % of the headers were selected by all subjects so most subjects share the same understanding of what constitutes a header. This indicates the possibility of new segmentation technique based on human understanding.

References

1. Web accessibility initiative, World Wide Web Consortium: Web Content Accessibility Guidelines 1.0, Web Content Accessibility Guidelines Website (1999) http://www.w3.org/TR/WCAG10/
2. Japanese Industrial Standard Committee: Japanese Industrial Standards X8341-3: Guidelines for older persons and persons with disabilities – information and communications equipment, software and services – Part 3: Web content. Japanese Standards Association, Tokyo (2004)
3. Watanabe, M., Okano, A., Asano, Y.: Universal Design Guidelines for Web Contents. NTT Technical Review 3(11), 17–22 (2005)
4. Asano, Y., Watanabe, M., Hamano, T., Ogawa, K.: Web Accessibility Guidelines that Solve Problems Peculiar to Japanese. In: Proc. WWCS 2004 pp. 601–606 (2004)
5. Asakawa, C., Itoh, T.: User Interface of a Home Page Reader. In: Proc. ASSETS 1998 pp. 149–156 (1998)
6. Lee, A.: Scaffolding Visually Cluttered Web Pages to Facilitate Accessibility. In: Proc. of the working conference on Advanced visual interfaces, pp. 90–93 (2004)
7. Yu, S., Cai, D., Wen, J., Ma, W.: Improving Pseudo-Relevance Feedback in Web Information Retrieval Using Web Page Segmentation. In: Proc. WWW 2003 pp. 11–18 (2003)
8. Krüpl, B., Herzog, M.: Visually guided bottom-up table detection and segmentation in web documents. In: Proc. WWW 2006 pp. 933–934 (2006)
9. Shaikh, A., Lenz, K.: Where's the Search? Re-examining User Expectations of Web Objects, Usability News, vol. 8(1) (2006), http://psychology.wichita.edu/surl/

Editing a Test Suite for Accessibility of Interactive Web Sites

Gerhard Weber[1] and Kurt Weimann[2]

[1] Technical University of Dresden, Mommsenstr. 13, 01062 Dresden, Germany
g.weber@hci-research.de
[2] University of Kiel - Multimedia Campus , Boschstr. 5,
24118 Kiel, Germany
k.weimann@hci-research.de

Abstract. A test suite is a set of tools and data to ensure quality of a software application such as a dynamic web site. Reliability of page by page web accessibility tests is difficult to achieve as this requires defining appropriate partitions of the set of all web pages. We propose a partitioning method based on additional information about reading behaviour by users especially for measuring readability and ease of navigation. An upper limit for the path length and hence size of the partition is developed. We present a test case editor to describe user group specific testing scenarios and additional types of tests.

Keywords: evaluation, software testing, web accessibility.

1 Introduction

Web accessibility is a quality feature of the web server, the web browser and AT requiring testing. Automatic testing can take all these modules into account while manual evaluation of web accessibility focuses on the information presented by the web browser. Testing of any software applications is a semi-automatic process consisting of several steps to ensure the application meets all requirements as specified before starting the software development. For web applications focus is commonly on functional and stress testing as browsers behave differently and the client/server system or multi-tier system may degrade its functionality at a certain level of complexity of message processing by the underlying message passing system. Functional testing is essential for each newly integrated software module, especially if changes were made due to usability requirements.

Figure 1 describes testing of a software application. First step is the design of test cases containing information about the result of some data processing if the software application is correct. Second step is the preparation of test data serving as input to the software application to be tested.

Next step requires running the software application with all test data of this particular test case. The result of this processing is then to be compared with the test case. A test report details discrepancies or correctness of the result.

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 193–201, 2007.
© Springer-Verlag Berlin Heidelberg 2007

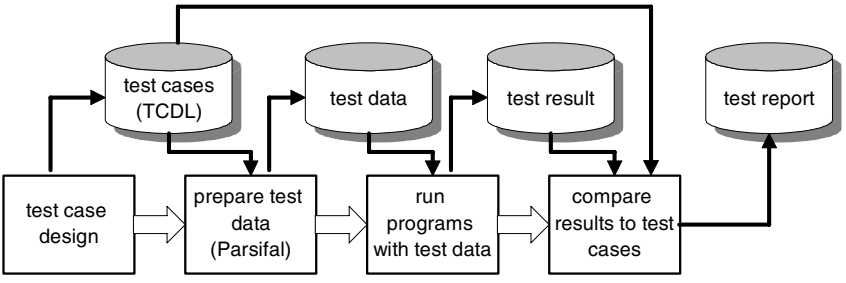


Fig. 1. Testing software applications

For black-box testing and structural testing [10] it is important to develop test cases which partition the set of all possible test data. Smallest partition is consisting of a single element, a particular web page with more or less mark-up data. A partition is called an equivalence partition if it is an equivalence relation[1]. By developing an equivalence partition it is only required to develop test cases for at least one element of the equivalence partition. As before, a partition consisting of a single element is an equivalence relation, however, it is not desirable to start testing for each and every web page the user agent (and the server) over and over again. Moreover, sometimes it is even difficult to compute one and the same web page again.

Our objective within EU-project BenToWeb is to validate by software testing if a particular equivalence partition can be developed, which covers accessibility requirements. Applying such a partitioning to arbitrary web pages will allow testing accessibility of web pages.

The paper is structured as follows. Section 2 discusses the development of mark-p language for testing accessibility. Section 3 investigates two critical accessibility requirements, describes differences among manual and automatic testing and introduces the concept of extended test cases. By developing partitions in sections 4, extensions can be restricted in their path length. Section 5 presents Parsifal as a tool to develop complex test cases.

2 Test Case Development for HTML

Test cases for testing web accessibility can be described by the Test Case Description Mark-up Language (TCDL) in XML [9]. A test case refers to a test file (web page) which may itself include mark-up for interaction techniques. Test cases are stored as TCDL files and support individualization of user agents including enhancements provided by assistive technologies (AT) like screen readers and screen magnifiers. TCDL files specify through user-group specific scenarios the type of user agent and AT needed for allowing the test files to comply with accessibility requirements. For example, does a screen reader read out the alt-Attribute in HTML through a speech synthesizer or presents it in Braille on a Braille display to a blind user. The result of testing is essentially either pass or fail. Detailed explanations can be more helpful for

[1] An equivalence relation is a set of pairs which are reflexive, symmetric and transitive.

fixing the inaccessible test file and require mark-up such as the Evaluation and Report Language [8]. Both, *pass* and *fail* are options essentially required in the sense that the failing test case identifies inaccessible web pages. For example, if an alternative description is inappropriately listing only a file name, or is consisting of white space, the test for accessibility is intended to fail.

3 Developing a Test Suite

A test suite is a set of test cases describing each a test file (web page) and test case data which may be extended for purposes of testing with user agents. A test suite is made up of passing and failing test cases in order to support automatic, manual and user testing of web pages.

Automatic testing is referring to a software testing tool checking web accessibility by inspecting mark-up of single test files [4] or multiple test files [11, 12]. We consider integration of such test tools as an important aspect to develop test cases consisting of arbitrary web pages. Automatic testing delivers a test report which identifies inaccessible mark-up of a web page.

Manual testing has to be performed by a human expert and not by an end-user, for example, if there is no appropriate AT available yet. More often manual testing is applied, when no algorithmic approach has been developed such as for testing simplicity of language. Manual testing is based on inspection of the test files and very often limited to a subset of web pages. For example, in Germany asks the popular BITV-Test to inspect three selected web pages[2] per web site.

User testing of test files refers to the conventional empirical approach requiring users to complete a task while using a web browser and navigating through hypertext. The result is either successful or unsuccessful task completion. The probability for achieving the task and its standard deviation are user group specific [1]. Users test accessibility by generating input to browse one and more commonly multiple test files by using their AT. User testing has to be repeated with the same task by several users with different user agents (browsers and AT) in order to be able to generalize the test result.

Figure 2 describes the scenario necessary to test web sites manually or automatically with a test suite applying also user input as test data. For testing accessibility of navigation within a page or to subsequent pages it is important to replicate the user input and controlling the browser with identical keystrokes and mouse movements. Technically, this requires in an initial step to record one user's input and playback the sequence of events for testing afterwards. Input both for the web browser and for controlling the AT have to be recorded.

Table 1 lists several types of automatic test tools for accessible mark-up features, an example, identifies some common limitations, and lists probabilities for an accessible test file. The purpose of the test suite is to identify if such limitations are encountered.

While the probability for validity of a particular accessibility barrier simply is p_i, some mathematical operator may be applied to integrate results for individual web

[2] http://www.bik-online.info/verfahren/bitvtest/kurzinfo.php

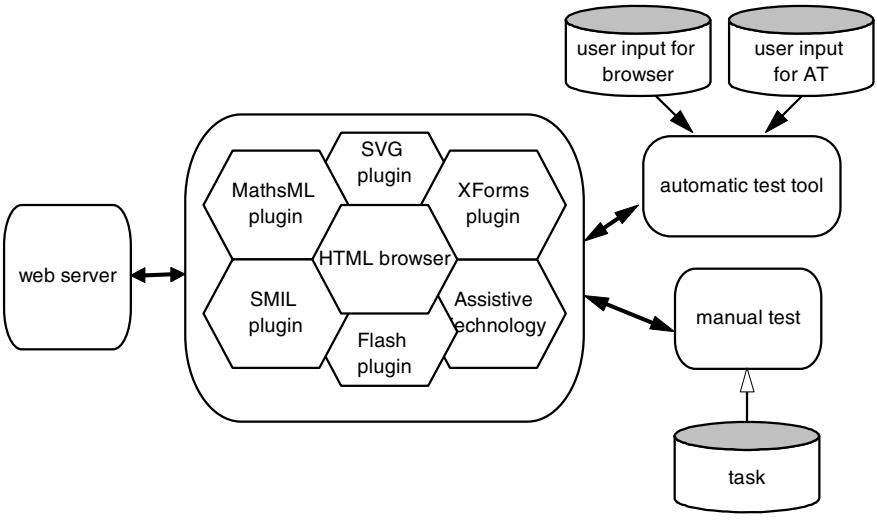


Fig. 2. Running automatic accessibility test tools

pages [17, 15]. The probability of a web site being accessible is a product of the probability of each individual web page's accessibility.

Table 1 refers to current challenges in automatic testing such as Guideline 14 "Ensure that documents are clear and simple" and Guideline 13 "Provide clear navigation mechanisms" of the Web Content Accessibility Guidelines 1.0 [13]. WCAG 2.0 [14] asks for similar requirements in Guideline 3.1 "Make text content readable and understandable" and Guideline 3.2 "Make the placement and functionality of content predictable" which is supported by the Level 2 Success Criteria 3.2.4 "Components that have the same functionality within a set of Web units are identified consistently".

Measuring complexity of natural language requires a comparison of linguistic features of natural language with those mastered by target audience of a web site. By developing a semantic representation of text such automatic comparison becomes feasible [5]. The criterion is passed if for example a dyslexic user with a university degree can read a scientific publication in her field. Measuring simplicity of language requires readability scales whose norms are based on mainstream school education. Readability scales are updated over time as well (typically over decades) and make up one element of baseline information upon which the test case has been developed. Baseline information hence not only has to describe the technological foundations but also the scales and measurement methods when applying accessibility criteria.

Likewise is "clarity of navigation" [2] difficult to test and document. It requires tracking visitors navigating a web site, learn their paths automatically and consistently transform this path into another path suitable to be used by AT. Detecting automatically whether a navigation path is followed can be of non-polynominal complexity [7] and probabilistic approaches can achieve polynominal complexity [6]. From a user's point of view navigation is supported by the AT. AT has its own

Table 1. Limitations and results of different types of test tools

type of test tool	**example**	**Limitations**	**probability of accessibility**
automatic tests of the validity of the mark-up language	XML parser, web browser, screen reader	contradictory use of multiple mark-up languages, e.g. <img> in SMIL and HTML	p_i
automatic tests checking particular accessibility features	colour contrast checking tool, acronym tester, mark-up test of foreign languages	changes in background colour in adjacent text regions	p_i
automatic tests checking multiple accessibility criteria of a web page	Hera[3]	ease of navigation, simple language	$f(p_i)$
automatic tests checking accessibility of complete web sites.	Lift Machine[4], WebXACT[5]	dynamic web pages being not replicable	$\prod_{i=1}^{k} f(p_i)$

shortcuts and hence must be documented together with a test case as well. In addition, many GUI-based screen readers for Braille output support routing [16], which creates artificial input to activate a link by a simulated mouse input or simulated regular keyboard input to support users for other reasons.

Not in any case are manual testing and automatic testing leading to the same result. If some automatic approach to test some feature is not yet available, the result is undefined. Manual testing, however, will be possible then, but requires developing of some task suitable for the web page or web site to be tested.

In order to overcome these limitations pertinent to the type of automatic test or inherent to the use of specialized scales or other data related to usability issues additional efforts for documenting the test case are required. Figure 3 describes these extended test cases.

[3] http://www.sidar.org
[4] http://www.usablenet.com
[5] http://webxact.watchfire.com/

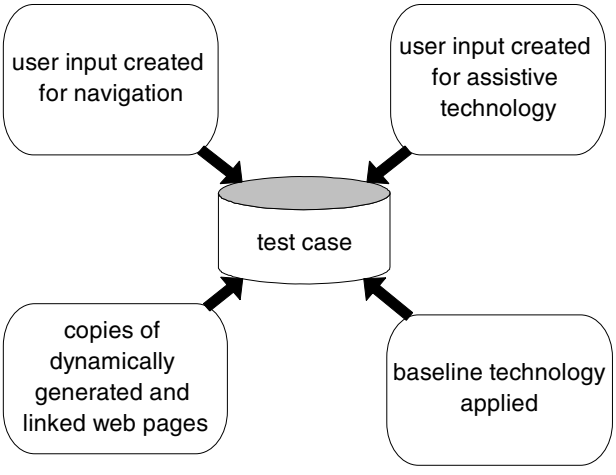


Fig. 3. Extended test cases

Any user input necessary for browsing the test files has be re-played in order to test either a set of web pages or a web page with links to bookmarks in itself. If pages are generated dynamically and cannot be replicated a copy must be made and, finally, test scales such as for readability have to be included as well.

4 Partitions

Extended test cases require many more test files compared to standard automated testing. Without any partitioning the computational resource can be considerable. For example navigation in an web site with auctions such as eBay may require copies of all test files linked to a particular search for a product. Even within a limited preview of a few days this may include thousands of files depending on the category.

A partitioning can be developed on the basis of Hypertext Probability Automata (HPA) and the probability linked web pages are accessible. A HPA denotes each web page as a state s_i and for the transition $q(s_i)$ a probability value [7]. Figure 4 illustrates an HPA.

The probability to reach from s_1 a particular s_k is

$$\prod_{i=1}^{k} p(s_i) \tag{1}$$

If

$$\prod_{i=1}^{k} f(p_i) > \prod_{i=1}^{k} p(s_i) \tag{2}$$

barriers are larger to overcome then the probability to reach a particular web page. All subsequent pages will not be reached due to inaccessibility. We can partition the test

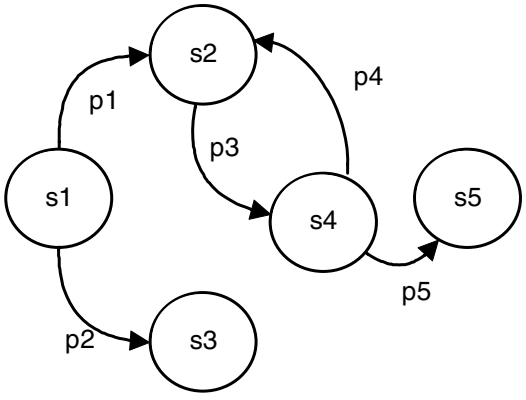


Fig. 4. Hypertext Probabilistic Automata

files accordingly and restrict the number of web pages to be tested. If the frequency of links visited cannot be obtained all paths leaving a web page still have to be tested.

5 Editor for Test Cases

We have developed an editor for editing test cases. The Parsifal tool is a graphical Test Case editor for test descriptions as defined by TCDL 1.1 specification [3]. Test

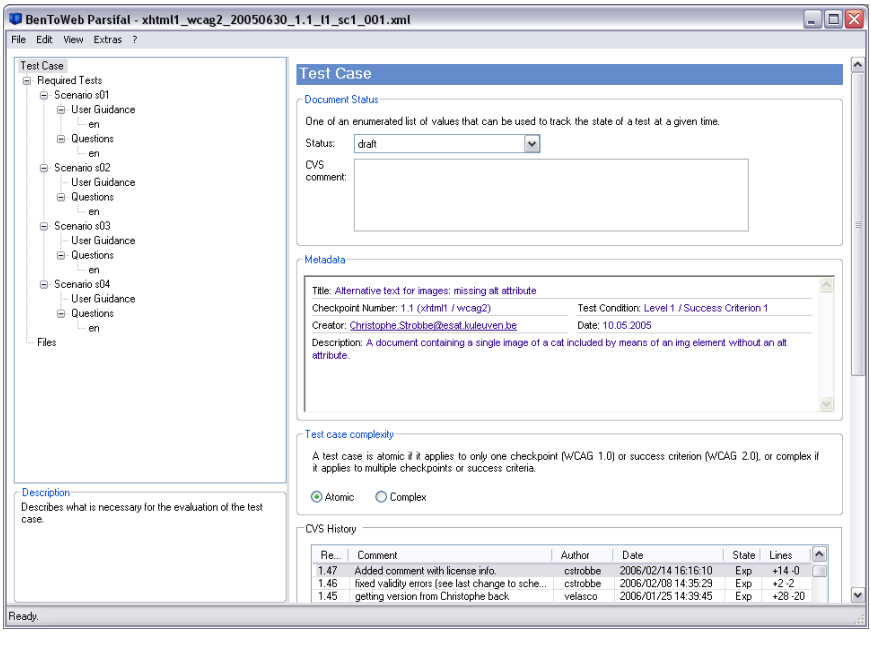


Fig. 5. Parsifal tool to edit test cases

case description files are written in XML. A graphical user interface allows test case editors to avoid XML coding. Parsifal is implemented in C# using .Net Framework 1.1. A setup routine for easy installation and runtime configuration is provided for Parsifal. In Parsifal parallel work on test case descriptions becomes possible by using a Concurrent Versions System (CVS). Figure 5 shows the GUI for a single test file.

Testing accessibility of multiple web pages becomes possible by linking the first test file with other test files. Automatic and manual testing can follow these links and apply navigation criteria. In TCDL such test cases are no more atomic but are called complex test cases and require testing for multiple accessibility criteria. Figure 6 shows how test case authors distinguish between different test modes.

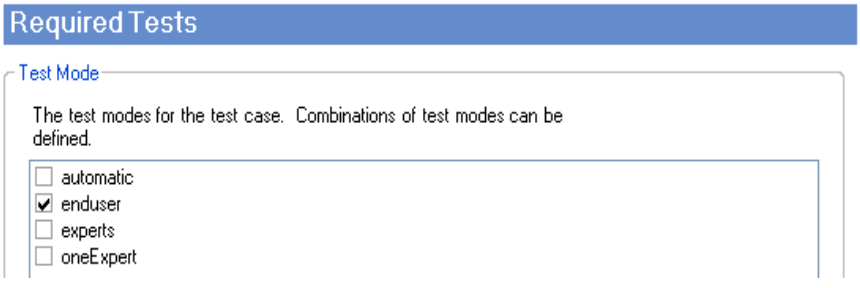


Fig. 6. Test modes supported by Parsifal

Interfacing with automatic test tools is achieved by publishing the test files through a web server. The automatic test tool follows the links. A macro recorder can record a user's input and the sequence of events in a file. Repeating manual tests becomes possible by playing back these events.

6 Conclusion and Future Work

Accessibility testing follows the same principles as any software testing if automatic testing can be performed. We have identified two challenges for automatic testing which require experts or users to perform a manual or user test. Both, simplicity of text for the intended readers, and predictability of navigation need more test data then just a web page. Measuring readability of text requires knowledge about the intended users. If probabilities of links visited are determined through manual evaluations, automatic tests can be ended when accessibility barriers are too high. The Parsifal test case editor allows specification of intended user groups through scenarios and enables test case authors to specify the type of test.

Future work will be needed to integrate user input with automatic tests of navigation barriers. It will be attempted to obtain probabilities of links visited from manuals tests and replaying a synthesized set of input events representing a particular user group.

Acknowledgements. This work has been undertaken in the framework of the project BenToWeb IST-2-004275-STP funded by the IST Programme of the European Commission.

References

1. Disability Rights Commission: The Web - access and inclusion for disabled people, The Stationary Office, London (2004)
2. Martin, G., Marcos, G., Rivera, C., Pintado, I., Vidau, A.: Testing Web Navigation for all: an Agent-based Approach. In: Miesenberger, K., et al. (eds.) ICCHP 2006. LNCS, vol. 4061, pp. 223–229. Springer, Heidelberg (2006)
3. Herramhoff, S., Petrie, H., Strobbe, C., Vlachogiannis, E., Weimann, K., Weber, G., Velasco, C.A.: Test Case Management Tools for Accessibility Testing. In: Miesenberger, K., et al. (ed.) ICCHP 2006. LNCS, vol. 4061, pp. 215–222. Springer, Heidelberg (2006)
4. Ivory, Y.M.: Automated Web Site Evaluation. Springer, Heidelberg (2004)
5. Jenge, C., Hartrumpf, S., Helbig, H., Osswald, R.: Automatic control of simple language in web pages. In: Miesenberger, K., et al. (eds.) ICCHP 2006. LNCS, vol. 4061, pp. 207–214. Springer, Heidelberg (2006)
6. Kukulenz, D., Pauli, J.: Prediction of navigation profiles in a distributed Internet environment through learning of graph distributions. Int. Journal of Computational Intelligence and Applications 2(3), 303–312 (2002)
7. Levene, M., Loizou, G.: Navigation in Hypertext is easy only sometimes. SIAM Journal on Computing 29, 728–760 (1999)
8. McCathieNevile, C., Abou-Zahra, S.: Evaluation and report language (EARL) 1.0 Schema. W3C editor's working draft (2006) Available at http://www.w3.org/WAI/ER/EARL10/WD-EARL10-Schema-20060101
9. Strobbe, C.: The BenToWeb XHTML 1.0 Test Suite for the Web Content Accessibility Guidelines 2.0 - Last Call Working Draft. In: Stephanidis, C. (ed.) HCI International 2007. LNCS, vol. 4556, pp. 160–167. Springer, Heidelberg (2007)
10. Summerville, I.: Software Engineering, 8th edn. Addison-Wesley Longman, Amsterdam (2006)
11. Velasco, C.: imergo - Development of automated tools that support standards compliance and quality assurance methods for the production and maintenance of Web portals. In: Proc. XML Europe 2004 (2004) available at http://www.idealliance.org/papers/dx_xmle04/ papers/03-02-01/03-02-01.pdf
12. Velasco, C.: Development of automatic Web accessibility checking modules for advanced quality assurance tools, this proceedings (2007)
13. W3C Recommendation. Web Content Accessibility Guidelines 1.0 (WCAG 1.0). Chisholm, W., Vanderheiden, G., Jacobs, I.(eds.): World Wide Web Consortium (MIT, INRIA, Keio) (May 5, 1999) available at http://www.w3.org/TR/WCAG10/
14. W3C Working Draft. Web Content Accessibility Guidelines 2.0 (WCAG 2.0). In: Caldwell, B., Chisholm, W., Vanderheiden, G., White, J. (eds.): World Wide Web Consortium (MIT, ERCIM, Keio) (2006) available at http://www.w3.org/TR/WCAG20/
15. Web Accessibility Benchmarking Cluster: D-WAB2 unified web evaluation methodology (UWEM 1.0) (2006) available from http://www.wabcluster.org/uwem/
16. Weber, G.: Adapting direct manipulation for blind users. In: Stacey, A., et al. (eds.) Adjunct Proceedings of INTERCHI '93, pp. 21–22. Addison Wesley, London, UK (1993)
17. Zeng, X.: Evaluation and Enhancement of Web Content Accessibility for Persons with Disabilities. PhD thesis, University of Pittsburgh (2004)

Discrimination and Perception of the Acoustic Rendition of Texts by Blind People

Vassilis Argyropoulos[1], Konstantinos Papadopoulos[2], Georgios Kouroupetroglou[3], Gerasimos Xydas[3], and Philippos Katsoulis[3]

[1] University of Thessaly, School of Humanities, Department of Special Education, Argonafton & Filellinon St, Volos, GR-38221, Greece,
vassargi@uth.gr
[2] University of Macedonia, Department of Education and Social Policy,
kpapado@uom.gr
[3] University of Athens, Department of Informatics and Telecommunications, Panepistimiopolis, GR-15784, Athens, Greece
{koupe,gxydas,}@di.uoa.gr

Abstract. This paper reports on the results from a series of psychoacoustic experiments in the field of the auditory representation of texts via synthetic speech which comprise similar acoustic patterns so called "paronyms". The errors which occur when listening to paronyms are classified as errors of phonological type. Thirty blind and thirty sighted students participated in psychoacoustic experiments. The results from the experiments depicted the types of the subjects' errors and addressed comparisons between the performances of blind and sighted students on their auditory distinctions towards the chosen scripts (paronym words and sentences with paronyms). The discussion considered the practical implications of the findings such as issues regarding education as well as the development of suitable design of acoustic rendition of texts in favor of better perception and comprehension.

Keywords: text-to-speech, synthetic speech perception, discrimination of synthetic speech, paronym words.

1 Introduction

Related works in the field of the acoustic representation of documents and texts (text-to-speech and document-to-speech) have pointed out the need for rendering with distinction peculiar linguistic patterns – such as paronym words - and part of the meta-information [1], [2], [3], [4].

More specifically, paronym words are characterized as linguistic structures which their acoustic representation bears a great resemblance between them and as such they might create misunderstandings or confusions to listeners. For example, the words "affect" → [af`ekt] and "effect" → [ef`ekt] in English and the words "φίλη" → [f`ili] and "φυλή" → [fil`i] in Greek are considered as paronyms. There are two reasons behind these misunderstandings when listening to paronyms. The first reason originates from the similarity of the phonemic representation of paronymal words. For

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 205–213, 2007.
© Springer-Verlag Berlin Heidelberg 2007

example: “έκλειψη” → [`eklipsi] and “έκθλιψη” → [`ekTlipsi]. The acoustic realization of these words can be confusing for the listener, since they only differ in the unvoiced fricative [T] that might not be heard (when saying [`ekTlipsi] or might be thought as being heard (when saying [`eklipsi]). Speech rhythm is a significant factor that affects this perception.

The second reason of the misunderstandings originates from the intonational realization of the phrases. There are cases where though the phonemic representations of paronyms are identical, their lexical stress placement differs. For example: “φίλη” → [f`ili] and “φυλή” → [fil`i]. According to the intonational theory [24], lexical stress has an impact on the perceived tonal stress however it is not the only factor that affects the position of the prosodic pitch accent in an intonational phrase. There are stronger events, such as the phrase accent, which usually define a different tonal prominence in the word/phrase than the one imposed by the lexical stresses themselves. Phrase accent is defined from a series of linguistic attributes, such as the phrase type (e.g. affirmative, negative, question, wh-question etc). As a result, written lexical stress does not always produce tonal rises; falls are also common and this depends on both linguistic and para-linguistic phenomena, as well as the common ground between collocutors. This normal inconsistency between the written lexical stress and the acoustic pitch accent alignment leads to acoustic misunderstandings between words with identical phonemic representation but different lexical stress placement.

Though paronym perception accuracy has not been studied in depth, there are works that study the intelligibility of segments in both natural and synthetic voices, under normal or noisy conditions. In [5], the error rate in phoneme discrimination was measured to be 2.4% in natural speech, which was raised to 9.0% when using synthetic speech (Festival TTS, diphone-based voice). Most of the errors occurred in the final consonant. In that work it was also shown that in reverberation environments the phoneme discrimination is even lower. These values are significant higher than older works which showed that phoneme error rate in intelligibility tests using rule-based synthetic voices was ranging from 12.92% to 73.97% [6]. Finally in [7], it was shown that the initial segments of CV and VC syllables are harder to identify than the final segments, independently of the noise level.

The aims of the study presented here were: i. to categorize the types of errors the blind and sighted students made when rendering paronym words ii. to evaluate blind and sighted students’ performances separately (within groups) and iii. to compare blind and sighted students’ auditory discriminations (between groups).

2 Method

Sixty (60) university students took part in this study; thirty (30) of them had severe visual impairments (21 totally blind and 9 partially sighted) and the other thirty (30) students were sighted. The age range for the group of the sighted students was from 21 years to 24 years (mean= 22.27, SD= 1.081) and the age range for the group of the blind students was from 18 years to 35 years (mean= 26.23, SD= 4.804). They were all students of different schools in Greek universities and apart from blindness had no other additional disabilities.

Each student was interviewed and audiotaped in a 60 minute session. The experimental design comprised two parts: structured interviews and experiments. The interviews – apart from the participants' personal details - were focused on issues such as the experience of the students in using assistive technology mostly in the usage of screen readers. Each item in the questions was rated on a 4-point Likert-type scale, from 1 ("I use it a lot) to 4 ("I don't use it at all").

After the interview, every student was invited to listen to structured scripts and was asked to give back what he/she heard from the aural presentation of the scripts. The scripts were divided into two sets. The first one comprised 12 pairs of paronym words (Test 1) and the second one comprised 17 sentences in which were included also paronym words (Test 2). The auditory scripts were composed in conjunction with appropriate modifications and non-speech insertions. The paronym words as well as the structure of the sentences were chosen from the research team (authors) after many internal tests. The hypotheses of this study relied on the performances of the blind and sighted students when doing the tasks in recognizing the "paronyms" in specific auditory scripts.

These experimental sessions were carried out by utilizing a synthetic voice instead of a natural one. The purpose behind this decision was the fact that text-to-speech synthesizers have become essential tools for the accessibility of blinds and modern human machine interaction applications. Thus, the nature of the experiments presented in this work originates from real operational conditions. Before running the experiments, all students listened to examples of synthetic speech generated by DEMOSTHeNES language platform [8], [9], [10]. All subjects – one by one – were given 5 minutes to familiarize themselves by listening to some examples of synthetic speech which their content was different from the content of the stimulus material (words & sentences). After the familiarization stage the stimulus material was presented aurally by DEMOSTHeNES to the students and the latter were invited to render what they heard. The participants could stop the acoustic rendition simply by pressing a button in a device and write down in their own pace what they were listening from the headphones. The only restriction which took place was the fact they could not reverse the acoustic rendition to elucidate a misunderstanding. There were approximately 5 seconds break between the acoustic representations of the pairs of the paronym words as well as between the sentences which included also paronym words.

The default synthetic voice that we used featured a well established corpus-based prosodic model [11] and the Mbrola synthesizer [12] with the Greek diphone database gr2 [8]. The prosodic baselines used were: pitch=110Hz, speed=140 words per minute and volume=100.

The whole procedure was driven by the researchers, and, in the end, all the students' answers were audiotaped, transcribed, organized, reviewed for errors, and analyzed using SPSS 14.0.

Paronym words are characterized as linguistic structures which their acoustic representation bears a great resemblance between them and as such they might create misunderstandings or confusions. For this the categorisation of the students' errors regarding paronyms was based on phoneme errors. The phoneme error pattern used in this study was a synthesis of other similar patterns [13], [14] and comprised seven categories which referred to errors of phonological type. Phonological type errors are

those that change the acoustic image of the word. These categories represent different cases such as omission of phonemes, addition of unneeded phonemes, wrong accentuation or different combinations of the above (see Table 1 for a list of all categories).

Table 1. Categories of Phonological-type errors (PTE)

Categories	Phonological-type errors
A	Accentuation (accent)
B	Phoneme substitution (first part of the word)
C	Phoneme substitution (middle and last part of the word)
D	Addition of a phoneme (first part of the word)
E	Addition of a phoneme (middle and last part of the word)
F	Combination of the following: i. Omission of more than one phoneme in the word ii. Addition of more than one phoneme in the word iii. Wrong rendering regarding accentuation combined with phoneme substitutions or omissions
G	Omission of the whole word or rendering of a different word

Category A referred to accentuation issues. Accent in Modern Greek is a short line which is written only above vowels indicating the way they are pronounced in the word.

The analysis below focuses on correlations between the groups of the participants (sighted and blind) and the categories of phonological-type errors (PTE) in two sets of psychoacoustic experiments: a. phonological-type errors in words (PTEW) and b. phonological-type errors in sentences (PTES).

3 Results

Preliminary analyses showed that according to our data a violation of the assumptions of normality took place; therefore, it was decided to use non-parametric tests. Also it is important to mention that the only mistakes the participants made were when they tried to render the paronym words. Hence it can be argued that the specific text-to-speech system did not constitute an external uncontrolled variable to threaten the validity of the experiments.

It was found from the interviews that the correlations between PTE and sex, age, degree of visual loss and frequency of usage of screen readers respectively were not significant.

It was also found that within the group of the participants with severe visual impairments there was not significant correlation between the number of PTEW and the number PTES. On the other hand, there was a medium positive correlation between the same variables within the group of the sighted students ($\mathbf{r_{spearman}}$ (2-tailed) = 0.368, $p<0.05$).

The Mann-Whitney U Test was conducted to compare the number of phonological-type error (PTE) made by the two groups (sighted & blind). The analysis indicated that there was a significant difference between the performances of the blind and the sighted participants regarding paronym words either in pairs or in sentences (U= 259.00, Z= -2.835, $p<0.01$). More specifically the number of the blind participants' PTE was significantly less than the corresponding number of the sighted ones. A more detailed analysis revealed that the differences between the two groups of the participants were significantly different only regarding the PTEW (U= 237.00, Z= -3.172, $p<0.005$) and not the PTES.

It was also attempted a more discernible analysis regarding the categories mentioned in Table 1. More analytically, Table 2 provides descriptive data for the blind students (see below).

Table 2. Means of PTEW & PTES in all categories of PTE regarding blind students

	PTEW		PTES		PTE	
	Mean	SD	Mean	SD	Mean	SD
Category						
A	0.97	1.426	1.27	1.484	2.23	2.269
B	0.23	0.430	0.00		0.23	0.430
C	0.40	0.724	0.07	0.254	0.47	0.860
D	0.27	0.521	0.00		0.27	0.521
E	0.57	0.568	0.10	0.305	0.67	0.711
F	0.40	0.675	0.37	0.556	0.77	0.858
G	0.70	1.149	0.10	0.403	0.80	1.157
Total	3.53	2.649	1.90	1.807	**5.43**	3.579

Table 3 (see below) also provides descriptive data for the sighted students' performances. The two tables indicated that blind students performed more accurately in recognizing paronym words in both tests (M=5.43) compare with the performances of their sighted peers (M=8.00).

Table 3. Means of PTEW & PTES in all categories of PTE regarding sighted students

	PTEW		PTES		PTE	
	Mean	SD	Mean	SD	Mean	SD
Category						
A	0.83	1.053	0.93	0.944	1.77	1.406
B	0.13	0.346	0.03	0.183	0.17	0.379
C	0.50	0.572	0.03	0.183	0.53	0.629
D	0.47	0.507	0.13	0.346	0.60	0.563
E	0.73	0.691	0.00		0.73	0.691
F	1.20	0.961	0.47	0.730	1.67	1.295
G	1.63	1.650	0.90	0.759	2.53	1.925
Total	5.50	2.113	2.50	1.548	**8.00**	3.040

The Mann-Whitney U Test showed that the two groups differed statistically in categories D, F and G. That is to say, that the number of the PTE were statistically significant more on behalf of the sighted students rather than for the blind in the following categories: i. D (U= 305.00, Z= -2.501, $p<0.05$), ii. F (U= 268.500, Z= -2.798, $p<0.01$) and iii. G (U= 190.00, Z= -3.971, $p<0.001$).

Finally, distinguishing the PTE into PTEW & PTES the Mann-Whitney U Test indicated that regarding the test with the pairs of the paronym words (PTEW) the sighted participants made significant more errors than their blind peers in categories F (U= 228.00, Z= -3.541, $p<0.001$) and G (U= 289.50, Z= -2.538, $p<0.05$), whereas, regarding the test with the sentences (PTES) the sighted students made significant more errors than their blind peers in categories D (U= 390.00, Z= -2.053, $p<0.05$) and G (U= 183.00, Z= -4.617, $p<0.001$).

4 Discussion

In general, the blind students' performances were at higher level of distinction regarding the number of errors (PTEW & PTES) in both set of scripts. The fact that blind students surpassed their sighted peers it might be happened due to the experience of the former in listening to a big variety of pre-recorded study materials when developing their literacy skills [15], [16], [17].

As mentioned in the Results section, there was a medium positive correlation between the number of the PTEW and the number of the PTES regarding the performances of the sighted group. That is to say that medium level of inaccuracy

regarding PTEW is also associated with medium level of inaccuracy regarding PTES. In other words the sighted students showed an internal coherence regarding their acoustic distinction towards paronyms. It seemed that the rendering of the paronyms caused a sort of confusion in the sighted ones rather than in the blind participants.

Additionally, the previous inference is supported by the results of the Mann-Whitney U Test. The analysis of the data revealed that there was statistically significant difference in the PTE scores of blind and sighted students, particularly in the test with the words (pairs of paronym words-PTEW). It seemed that the sighted students faced difficulties when they were asked to identify and render the pairs of the paronyms whereas they did not experience the same difficulties when they invited to do the same in the second experiment (paronym words embodied in sentences – PTES). Therefore, it might be argued that the second experiment with the sentences provided a frame of a logical content which eventually helped the sighted participants to figure out more easily the paronym words. Blind students in Greece especially those who attend university do not have the chance to read their books in Braille whereas their sighted peers do not have such problems. The Greek Ministry of Education does not provide students in tertiary education with books in Braille. As a result most university blind students are oriented to aural reading. Hence the acoustic images of the words – including the paronyms- might have been grouped in a more concrete and distinctive form in blind students' memory. The fact that blind students use more the aural way of reading is supported by the findings of many researchers by highlighting the phenomenon of the decline in using Braille [18], [19].

Also it is worth noting that the classification system of the phonological type errors adopted in this study, seemed to offer the opportunity for a detailed analysis regarding the different kind of errors. In particular, Tables 1 & 2 presented the means and standard deviations of the performances of both groups (pair of words and sentences). The dispersion of the errors in the word tests was larger in the categories F and G in favor of the sighted students which in turn entailed a significant differentiation between them in terms of statistics with respect to the two student groups (see also Results). Category F refers to a combination of errors and it might also reflect the confusion the sighted students experienced when performing the tasks. Category G was found to be common in PTEW and in PTES by which the two student groups differed significantly. As mentioned in the Method section (see Table 1), category G refers to omission of the paronym word or to a substitution of the paronym word with another. It may be argued that the sighed students used more intellectual links to render the words or the sentences they heard during the tasks and as a result the degree of the consolidation of the paronym words as intellectual schemata [20] was decreased.

5 Conclusions

As mentioned in the introduction there is little research on issues germane to paronym acoustic realization and as such it was felt that this study is important in helping to fill that gap. The results of this research may be considered of great interest because they are strongly linked with educational issues. The practical implications of intonational theories have an important impact on the auditory representations of documents which

in turn may facilitate or hinder the education of blind and sighted students in schools [21], [22], [23].

The adopted categorization system regarding phonemic errors (see Table 1) seemed that served well this study and provided more insights about the nature and peculiarities of the errors made by blind and sighted students. Also these issues are strongly linked to spelling which in turn leads to the area of literacy and generally speaking to the curriculum [14].

Finally there is need for further research increasing the number of the participants and conducting experiments in different periods of their schooling. This data may lead to improvements of the synthetic speech regarding the distinctive rendering of special linguistic patterns such as paronym words in conjunction with the integration of phonemic error patterns.

Acknowledgements

The work described in this paper has been funded by the European Social Fund and National Resources under the HOMER project of the Competitiveness Programme: PENED, Greek General Secretariat of Research and Technology.

References

1. Raman, T.V.: An Audio View of (LA)TEX Documents, TUGboat. In: Proceedings of the 1992 Annual Meeting vol. 13(3) pp. 372–379 (1992)
2. Hakulinen, J., Turunen, M., Raiha, K.: The Use of Prosodic Features to Help Users Extract Information from Structured Elements in Spoken Dialogue Systems. In: Proceedings of ESCA Tutorial and Research Workshop on Dialogue and Prosody, Eindhoven, The Netherlands, pp. 65–70 (1999)
3. Shriver, S., Black, A., Rosenfeld, R.: Audio Signals in Speech Interfaces. In: Proceedings of International Conference on Spoken Language Processing, Beijing, China (2000)
4. Xydas, G., Argyropoulos, V., Karakosta, T., Kouroupetroglou, G.: An Experimental Approach in Recognizing Synthesized Auditory Components in a Non-Visual Interaction with Documents. In: Proceedings of the 11th International Conference on Human-Computer Interaction (HCII 2005) Las Vegas, Nevada SA, vol. 3, pp. 411–420 (2005)
5. Venkatagiri, S.H.: Segmental Intelligibility of Three Text-to-Speech Synthesis Methods in Reverberant Environments. Augmentative and Alternative Communication 20(3), 150–163 (2004)
6. Logan, J.S., Greene, B.G., Pisoni, D.B.: Segmental intelligibility of synthetic speech produced by rule, J. Acoust. Soc. Am. 86, 566–581 (1989)
7. Cutler, A., Weber, A., Smits, R., Cooper, N.: Patterns of English phoneme confusions by native and non-native listeners, J. Acoust. Soc. Am. 116(6), 3668–3678 (2004)
8. Xydas, G., Kouroupetroglou, G.: The DEMOSTHeNES Speech Composer. In: Proceedings of the 4th ISCA Tutorial and Research Workshop on Speech Synthesis, Perthshire, Scotland, pp. 167–172 (2001)
9. Xydas, G., Kouroupetroglou, G.: Text-to-Speech Scripting Interface for Appropriate Vocalisation of e-Texts. In: proceedings of EUROSPEECH, Aalborg, Denmark pp. 2247–2250 (2001)

10. Xydas, G.: Machine Learning Models and Selection Methods for the Auditory Representation of Documents via Synthetic Speech with Enriched Prosody, PhD dissertation, Department of Informatics and Telecommunications, National and Kapodistrian University of Athens, Greece (2006)
11. Xydas, G., Spiliotopoulos, D., Kouroupetroglou, G.: Modeling Prosodic Structures in Linguistically Enriched Environments. In: Sojka, P., Kopeček, I., Pala, K. (eds.) TSD 2004. LNCS (LNAI), vol. 3206, pp. 521–528. Springer, Heidelberg (2004)
12. Dutoit, T.: An Introduction to Text-to-Speech Synthesis. Kluwer Academic Publishers, Boston, MA (1997)
13. Tindal, G.A., Marston, D.B.: Classroom-based Assessment: Evaluating Instructional Outcomes. Columbus, OH: Merrill (1990)
14. Argyropoulos, S.V., Martos, C.A.: Braille Literacy Skills: An Analysis of the Concept of Spelling. Journal of Visual Impairment & Blindness 100(11), 676–686 (2006)
15. Koenig, A.J., Holbrook, M.C.: Literacy Skills, In: Koening, A. J. Holbrook, M. C. (eds) Foundations of Education. vol. II, AFB Press, pp. 264–312 (2000)
16. Kapperman, G., Sticken, J.: Assistive Technology. In: Koening, A. J., Holbrook, M. C. (eds.) Foundations of Education. vol. II AFB Press, pp. 500–516 (2000)
17. McCall, S.: Accessing the curriculum. In: Arter, C., Mason, L. H., McCall, S., Stone, S. (eds.) Children with visual impairment in mainstream settings. London: David Fulton, pp. 29–40 (1999)
18. Miller, B.: Spelling Bees and Grammar Gorillas. The Braille Monitor (November 1999) vol. 42(9) (Accessed August 8, 2004) Available online at http://www.nfb.org/slate/slss9904.htm
19. Argyropoulos, V., Martos, A., Leotskakos, B.: Blind students and spelling: An investigation into braille literacy skills. In: Proceedings of the ICEVI European Conference 2005: Education – aiming for Excellence, Chemnitz (2005) pp. 180–185 (2005)
20. Wadsworth, B.J.: Piaget's theory of cognitive and affective development (4th edn.). New York: Longman (1989)
21. Amato, S.: Standards for competence in Braille literacy skills in teacher preparation programs. Journal of Visual Impairment and Blindness 96(3), 143–154 (2002)
22. Arter, C., Layton, L.: Reading Preferences of Pupils with Visual Impairment. The. British journal of Visual Impairment 18(1), 41–44 (2000)
23. Fellenius, K.: Computer-based instruction for young braille readers in mainstream education-An evaluation study. Visual Impairment Research 1, 147–164 (1999)
24. Ladd, D.R.: Intonational phrasing: The case for recursive prosodic structure. Phonology 3, 311–340 (1986)

Using Speech Recognition and Intelligent Search Tools to Enhance Information Accessibility

Keith Bain[1], Jason Hines[2], Pawan Lingras[2], and Yumei Qin[3]

[1] Liberated Learning Consortium, Saint Mary's University, Halifax, NS B3H 3C3, Canada,
keith.bain@stmarys.ca
[2] Mathematics and Computer Science Department, Saint Mary's University, Halifax, NS B3H 3C3, Canada,
j_hines@cs.smu.ca
3IBM China Human Ability and Accessibility Center, 25/F, IBM Tower, Pacific Century Place, 2A Gong Ti Bei Road, Chaoyang District, Beijing 100027, PRC

Abstract. Access to spoken information presents challenges in educational, work, and public settings for many individuals. Speech Recognition technology offers a potential strategy for increasing access in various settings, including the university lecture environment. An international research team has successfully experimented with providing speech recognition generated lecture notes, available online as multimedia enhanced transcripts. Additionally, a specially designed search and indexing tool was introduced to enhance the user experience. An exploratory study was conducted to gauge student reactions, identify usage patterns, and assess ease of use.

Keywords: speech recognition, transcription, multimedia, search, retrieval, indexing.

1 Introduction

Although physical access has improved in educational, employment, and public environments, information accessibility remains a significant barrier for many individuals. Information accessibility challenges abound in everyday situations. Blind and visually impaired persons face challenges accessing print information without intervening assistive technologies. Standard electronic documents can likewise introduce barriers. For example, early PDF documents were largely inaccessible for visually impaired users. Without available captions, digital audio is similarly inaccessible for those who are Deaf and/or Hard of Hearing. Other stakeholders, such as English Second Language learners, often struggle with audio information delivered in non-native languages.

Efforts to improve access to spoken information have garnered increased attention. For example, the Liberated Learning Consortium (www.liberatedlearing.com), comprised of fifteen member institutions in eight countries, is researching and developing new speech recognition (SR) technologies that improve information accessibility. Using SR to caption real-time speech and generate multimedia enhanced transcripts is subsequently referred to as the Liberated Learning concept. The initial

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 214–223, 2007.
© Springer-Verlag Berlin Heidelberg 2007

driver for Liberated Learning was dissatisfaction with conventional, intermediary based note taking approaches for students who were unable to write their own notes due to disabilities. Targeted stakeholders initially included those with physical and hearing disabilities, given these populations were most likely to require note taking support. The earliest studies considered the viability of using SR as live captioning and transcription tool for creating class notes in a university lecture. For these early software generated "notes", recognition errors were corrected, punctuation was added, formatting applied, summarization considered, and in some instances the prose itself was improved. These interventions were deemed necessary to transform the raw transcripts into a form more aligned with traditional definitions of class notes. These activities added significant overhead to the entire concept workflow and were generally acknowledged to be unsustainable given available resources. A gradual paradigm shift moved researchers away from the notion of providing SR based notes to the provision of SR generated transcripts. Activity therefore shifted from document enhancements to exploring various data alternatives, strategies for integrating archival and distribution platforms, and most importantly, enhancing user experience.

Two principal challenges emerged. Early users complained about the inherent sequential access nature of the text based transcripts, which were provided in hard copy for interested students. Because typical sixty minute lectures contained between 7,500 and 10,000 words, printing costs were prohibitive, as well as logistically difficult to manage. Transcripts were gradually made available through an online notes system. However, the majority of students downloaded the text documents to their local machines, where they typically assumed the burden of printing the entire lecture. Although students responded positively to the availability of SR generated transcripts, they reported an interest in accessing a variety of data formats which suited their individual learning preferences. They also expressed interest in the availability of explicit tools that would facilitate more efficient and intelligent utilization of the transcripts [1].

Two simultaneous efforts emerged to address these issues. First, developers designed a new SR application called IBM ViaScribe, which among other innovations, produced multimedia-based transcripts [2]. Using the W3C standard, developers published Synchronized Multimedia Integration Language (SMIL) based outputs. SMIL provided a standardized method for maintaining the synchronization between the resulting text transcription and the instructor's original audio, while in theory allowing students to access content through freely available media players. Researchers also published data using Microsoft's Synchronized Accessible Media Interchange (SAMI), given Window's Media Player did not conform to W3C standards. Eventually, multimedia output was integrated directly into an html structure.

Secondly, researchers began to conceptualize ways to improve machine readability of these multimedia transcripts and develop intelligent search and retrieval tools that could improve the user experience. A key premise was that students did not want to necessarily listen to/read/review the lecture in its entirety, but rather have control over the transcript experience. Researchers believed that students would opt to self select specific sections, search for key words, or extract concepts that were individually important for subsequent study. Furthermore, faculty users could potentially benefit

from tools that helped them revisit particular lectures or index keywords from a corpus of SR generated course transcripts.

2 Search and Retrieval of Keyphrases

Early information retrieval systems used Boolean logic to search for existence of the given keywords in a document [3]. Salton proposed the use of relative weights of terms in what he termed as the vector space model. In the conventional vector space method [4], the document collection is represented by an $m \times n$ term-document matrix, where *m* is the number of terms and *n* is the number of documents. Typically this matrix has fewer than 1% nonzero entries. Queries are represented as *m*-vectors, and a matrix-vector product produces an *n*-vector of scores that is used to rank the documents in relevance.

For convenience, initially, the vector space model (VSM) used automatically extracted keywords for identifying relevant documents. However, VSM can just as easily be applied to keyphrases. Previous research has shown that using keyphrases to assist the information retrieval processes can increase the retrieval efficiency.

Fagan [5] and Mitra [6] show that phrase-based indexing helps to increase the precision of the overall retrieval, especially at lower relevance ratings, even though it does not significantly improve the result at higher relevance ratings where phrases tend to overemphasize a particular aspect of the query. Anick [7] developed a relevance feedback system Paraphrase based on the observation that "lexical dispersion" of a term predicts the likelihood of the term being a topical term. The lexical dispersion is defined as the number of different noun compounds containing that term appearing in a given document set. Anick contended that noun compounds or keyphrases provide more discriminating power than individual terms. Phrases were extracted from the initial result set and presented to the user for relevance feedback. Another such attempt in the form of a document retrieval system called Phrasier was reported by Jones [8]. Phrasier ranked the similarity between documents by the degree of keyphrases overlapping in the documents. Here the query is also treated as a document. Results from human evaluation of Phrasier indicate that the phrase-based retrieval system is as effective as a full-text retrieval system while requiring less storage space for indexing. Buckland et al. [9] use noun phrases in the titles, authors and/or abstracts and metadata vocabularies such as classifications to build an "Entry Vocabulary Module" (EVM). EVM associates the noun phrases and the metadata vocabularies with the likelihood ratio statistics. The results are used as a dictionary for translating user queries in a natural language, such as English, to corresponding metadata vocabularies that are unfamiliar to the user. Since utilizing phrases is beneficial in many aspects in information retrieval, finding a set of keyphrases for a document repository that is suitable for both indexing and document retrieving is essential to improving efficiency of retrieval systems. Keyphrases also play an important role in exchanging and retrieving information with multiple repositories. In addition, keyphrases can help users get a better understanding of the content of a collection. They provide sensible entry points for the collection, show how queries can be extended, facilitate document skimming by visually emphasizing important phrases, and offer a powerful means of measuring document similarity [10] [11].

Traditionally, keyphrases are chosen manually. Authors assign keyphrases to documents they have written. Professional indexers often choose phrases from a predefined *controlled vocabulary* relevant to the domain at hand [12]. However, most of the documents on the web do not have keyphrases associated with them. Manual assignment of keyphrases is a tedious process that requires knowledge of the subject matter. Therefore, researchers are developing automatic extraction techniques.

There are two fundamentally different approaches to keyphrase generation: keyphrase assignment and keyphrase extraction [13]. Both use machine learning methods, and require a set of documents with associated keyphrases for training. Witten, *et al* [14] proposed Kea, keyphrase extraction algorithm. Kea generates candidate phrases by looking through the input document for any sequence of one, two, or three consecutive words. The consecutive words must not be separated by punctuation and must not begin or end with stop words. Stop words are commonly occurring words such as "the", "of", "to", "and", "he". Candidate phrases are normalized by converting them to lower case and stemming them. Stemming is a process of linguistic normalisation, in which the variant forms of a word are reduced to a common form, for example, the words: "connection", "connections", "connective", "connected", and "connecting" are all stemmed to "connect". Kea then uses the naive Bayes algorithm to learn to classify the candidate phrases as either keyphrase or nonkeyphrase. Turney [14] provided evidence that statistical word association can be used to improve the coherence of keyphrase extraction, resulting in higher quality keyphrases, measured by the degree of overlap with the authors' keyphrases. Furthermore, the new coherence features are not domain-specific. This work has been further extended to extraction of keyphrases from spoken language documents by Inkpen and Desilets [15] [16].

3 Note Finder Personal Search Engine

Utilizing the above keyword/keyphrase generation and extraction approaches, a prototype search and retrieval tool, named “Note Finder” was developed in 2005. Note Finder provides searching abilities based on keyphrases and stemming, automatic phrase table generation, listing of all the course documents, and the ability to crawl and index external websites containing various file types. The system updates automatically on a scheduled basis. Note Finder is built on the open source LAMP (Linux, Apache, MySQL, and PHP) framework which makes the source code accessible.

Note Finder works like Kea, with the exception that a phrase is always two words. Each indexed word has two lists associated with it; a same-stem list, consisting of words which have the same stem, and a proximity list, consisting of words which appear within a threshold distance of +/- five words (and do not contain any stop words).

The same-stem list is useful for displaying the various forms of a word in the phrase table. The proximity list consists of quadruplets (keyword, frequency, average distance, and minimum distance) which are used to determine the relevance of keyphrases. Using these two lists, a software agent automatically generates a phrase table (Fig. 1) each time the system is updated.

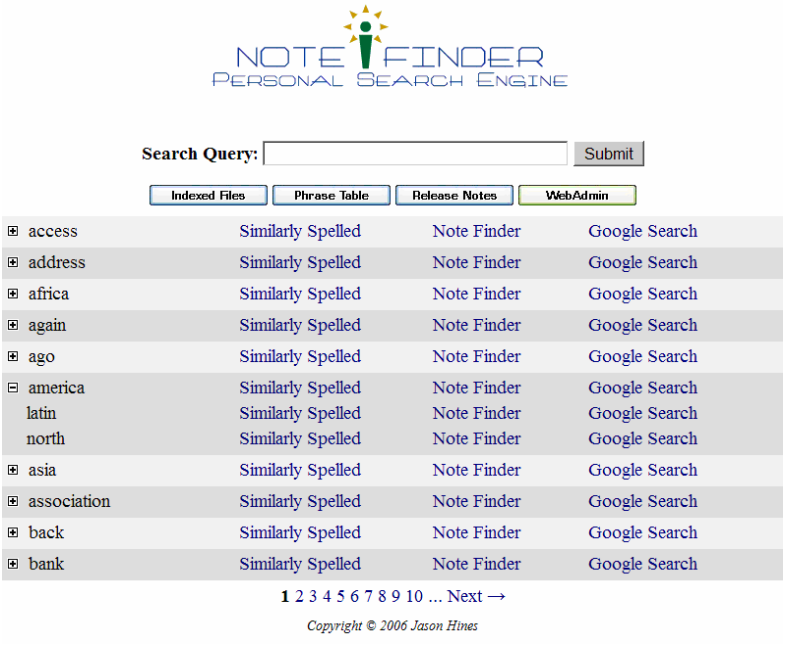


Fig. 1. Note Finder Phrase Table

Unlike conventional search engines, users have access to the top one hundred keyphrases. The phrase table gives users the ability to see the keyphrases which exist within the document set and also provides hyperlinks to quickly search for those keyphrases. This feature is especially useful for aiding memory recall. For example, users may be looking for a unique phrase but may not remember the exact semantics of the phrase. Since the phrase table presents users with keyphrases and a list of similarly spelled words, there is a greater chance that the user will find what they are looking for.

Fig. 1 shows the layout of the phrase table. From left to right, users are first presented with a collapsible menu of root keywords, a 'Similarly Spelled' link to display words which are similar to that keyword (contain the same stem), a 'Note Finder' link to search for that keyword in the Note Finder system, and a 'Google' link to search for that keyword using Google. By expanding the root keywords in the collapsible menu, users are presented with a sub-list of keywords. Each of the keywords in the sub-list belongs to a phrase containing the root keyword. As shown in Fig. 1, "latin america" and "north america" are keyphrases. Users can then search for these keyphrases by clicking on the corresponding 'Note Finder' or 'Google' link. The phrase table is sorted in alphabetical order and only displays ten keyphrases per page. To make the phrase table easier to read, different background shading is used for each alternating set of keyphrases.

Note Finder also provides a section for users to retrieve notes manually. This is useful for instances where users know exactly what document they are looking for.

This is all done in the 'Indexed Files' page of the system and is presented similar to that of the phrase table with the exception of using a collapsible directory structure.

Search results are quite similar to that of any major search engine but again Note Finder makes use of background shading to clearly differentiate between results. The document title is presented as a hyperlink through which the user accesses the current document. In instances where the document title is non-existent the actual filename is presented. If the document is a Liberated Learning enhanced multimedia document then users are presented with a link to the standard text format of the document as well as a link to the multimedia enhanced version of the document. Users also have the option to view a cached version of the document in which a list of same-stem search terms will be automatically highlighted.

4 Evaluating SR Multimedia Transcripts and Note Finder

In the 2006 fall semester, the first full term usage of Note Finder occurred. Since the system seemed robust from a technical and administrative perspective, the project team wanted to assess whether students found the system useful. Two, introductory undergraduate courses in Sociology and Anthropology were selected for evaluation. Both professors were experienced users of SR, having utilized the "Liberated Learning" approach for a number of years. The courses were also largely lecture based with large enrolments. In most classes, the professors used IBM ViaScribe to transcribe their lectures. The resulting multimedia transcriptions were saved to an automated network archival system [17]. A third party researcher subsequently uploaded the files to a dedicated server that hosted the Note Finder system. The Sociology professor additionally uploaded the multimedia transcripts to an internal network drive, which contained other course materials (syllabus, readings, etc). This drive was only accessible while physically on campus by logging into the university's local area network. The Anthropology lectures were only available through the Note Finder interface, which was accessible over any standard browser. The digitized lectures were typically available by the end of the day.

To help ensure feedback was as unbiased as possible, the researchers attempted to reinforce that system usage was not tied to course evaluation. At the beginning of the semester, an independent researcher provided both classes with an overview of the Liberated Learning concept and explicit instructions on accessing and using Note Finder. It was conveyed that the professors would not have system access and therefore would not be privy to usage statistics. Similarly, students were informed that researchers would not analyze usage on an individual basis. Throughout the term, both professors indicated they did intermittently remind students of the availability of the multimedia notes and Note Finder tool, but did not attempt to gather anecdotal feedback or discern if particular students were using the system.

For evaluation, a simple survey (Fig. 2) was developed and administered to gauge student reactions. There was an assumed implicit linkage between the SR generated transcripts and the tool students would use to access these transcripts. In other words, measuring reactions to the Note Finder tool itself would provide insights into the perceived underlying value of SR generated multimedia transcripts. There were no specific questions about the SR transcripts themselves (i.e. utility of available

formats). The paper based survey was attached to the end of each course's final examination, which was collected by an independent invigilator. The survey clearly indicated that participation was voluntary and would be treated anonymously. The professors were not present when the data was collected, nor did they have access to the results at any time.

1. Were you provided information on the use of the Note Finder as a tool for accessing lecture notes online?	**Yes**	**No**
If "Yes":		
2. Did you find the Note Finder easy to use?	**Yes**	**No**
3. Did you find the Note Finder helpful?	**Yes**	**No**
4. When using Note Finder, did you use (circle one):	• **Keyword search** • **Phrase Table** • **Both**	
5. Would you recommend Note Finder to other students?	**Yes**	**No**

Fig. 2. Student Survey

For the purposes of this initial analysis, surveys that were blank, incomplete, or contained a clear indication that Note Finder was not used, were not included. Out of 355 surveys collected, over seventy percent of the surveys were completed. Of those excluded, approximately 26% indicated the system was not used or were incomplete.

The results highlighted in Fig. 3 were extremely positive and consistent between both classes. As expected, the large majority of students found the tool intuitive and helpful. Most students reported using the keyword search or both Keyword and Phrase table. Only six percent reported using the phrase table in isolation. The most noteworthy class difference was the number of students that reported not using the

	Anthropology		*Sociology*		*Total*	*Average*
Surveys Analyzed	141		111		252	
Aware - Note Finder	134	95%	107	96%	241	**95.63%**
Easy to Use	126	89%	98	88%	224	**88.89%**
Helpful	120	85%	95	86%	215	**85.32%**
Recommend	126	89%	99	89%	225	**89.29%**
Keyword	68	48%	59	53%	127	**50.40%**
Phrase	9	6%	6	5%	15	**5.95%**
Both	63	45%	46	41%	109	**43.25%**

Fig. 3. Note Finder Survey Results

system (Sociology 16% versus Anthropology 9%). This difference is believed to be attributed to the availability of an alternative access point (LAN) for the Sociology students.

A number of surveys contained confusing contradictions. For example, six respondents did not find the Note Finder easy to use or helpful, yet indicated they would recommend it to other students. Conversely, five students indicated the tool was easy to use and helpful, yet said that they would not recommend Note Finder to their peers.

Many respondents included written comments about their experience. Most were brief and reinforced either the individual's like or dislike of the system:

> *A source for those who find it hard to study from the textbook or their own notes….an easy way to find in-depth notes…very helpful if you need to find something quickly and it provides thorough information.*

Some anecdotal feedback spoke to difficulties with network access, availability of other resources, or redundancy. Although these are variables beyond our control, they contributed to some negative responses:

> *I tried to get onto NoteFinder but couldn't get in….it asked for my A number* (student ID) *and password but I couldn't figure out which one it wanted…I didn't have any reason for using Note Finder given the notes were available on the P:Drive*

Although the overall results were encouraging, the methodology used afforded exploratory insights at best. The unsophisticated survey provides a 'panoramic' view of user satisfaction, but does not enlighten developers on specific system features that require redesign or improvement. For example, although students reported high usability, there is no data available to target areas for improvement.

Another question that remains unanswered is whether it was the Note Finder system itself that generated the positive responses, the underlying SR generated transcripts, or the combination of both. Our initial hypothesis is that both are desirable in unison. When previously presented with SR generated transcripts without any advanced search tools, students still responded positively, but provided suggestions for enhancing usability. The introduction of intelligent search and retrieval tools should therefore improve the user experience.

5 Further Work

The Note Finder system discussed here provides an excellent framework for further improving the student experience. In particular, we hope to enhance notes to include hyperlinks between notes based on indexed phrases. Whenever students are reading the notes and come across an unfamiliar phrase, they will be able to follow the link and get more information from other class notes. The current format of the notes can also enable us to identify and tag (using new XML tags) parts of the document as *doclets* that describe a particular concept. These tags will make it possible for us to

create a simple Semantic Web that can retrieve portions of the notes or "doclets" that are relevant to the user information needs, instead of retrieving the entire document.

Another interesting tangent is exploring social learning applications that can be facilitated through Note Finder. For example, students accessing SR transcripts could annotate sections they review, thus creating a searchable archive of peer transactions. A user could therefore see which sections were most frequently accessed by peers, perhaps indicating a socially generated measure of content importance.

6 Conclusion

Using SR to create multimedia transcripts available as "notes" for subsequent review seems to enhance learning opportunities in the university lecture environment. Coupled with intelligent search and retrieval tools like Note Finder, future generations of students could benefit from increased information accessibility. Determining whether the availability of SR generated multimedia transcripts and intuitive search tools improve academic performance is but one of many outstanding challenges facing researchers.

To tackle such questions, the Liberated Learning Consortium remains dedicated to continuing to investigate and advance SR applications that improve information accessibility. It is hoped that these efforts not only engender more sophisticated technologies, but also increase the necessary research collaborations that fuels such advancements.

References

1. Leitch, D., MacMillan, T.: Liberated Learning Initiative Innovative Technology and Inclusion: Current Issues and Future Directions for Liberated Learning Research. Saint Mary's University, Nova Scotia. Retrieved February 7, 2007 (2003), from http://www.liberatedlearning.com
2. Bain, K., Basson, S., Faisman, A., Kanevsky, D.: Accessibility, transcription, and access everywhere. IBM Systems Journal. vol. 44(3) (2005) http://www.research.ibm.com/journal/sj/443/bainaut.html
3. Salton, G.: Automatic Text Processing: The Transformation, Analysis, and Retrieval of Information by Computer. Addison-Wesley, London, UK (1989)
4. Salton, G., BUCKLEY, C.: Term weighting approaches in automatic text retrieval. Information Processing Management 24, 513–523 (1988)
5. Fagan, J.L.: The Effectiveness of a Nonsyntactic Approach to Automatic Phrase Indexing for Document Retrieval. Journal of the American Society for Information Science 40(2), 115–132 (1989)
6. Mitra, M., Buckley, C., Singhal, A., Cardie, C.: An Analysis of Statistical and Syntactic Phrases. In: Proceedings of RIAO97, Computer-Assisted Information Searching on the Internet, Montreal, Canada, pp. 200–214 (1997)
7. Anick, P.G.: The Paraphrase Search Assistant: Terminological Feedback for Iterative Information Seeking. In: Proceedings of SIGIR'99, pp. 153–161 (1999)
8. Jones, S.: Phrasier: a System for Interactive Document Retrieval Using Keyphrases. In: Proceedings of SIGIR'99, pp. 160–167 (1999)

9. Buckland, M.: Mapping Entry Vocabulary to Unfamiliar Metadata Vocabularies, D-Lib Magazine (1999) http://www.dlib.org/dlib/january99/buckland/01buckland.html
10. Gutwin, C., Paynter, G.W., Witten, I.H., Nevill-Manning, C.G., Frank, E.: Improving browsing in digital libraries with keyphrase indexes, Technical Report, Department of Computer Science, University of Saskatchewan, Canada (1998)
11. Witten, I.H., Paynter, G.W., Frank, E., Gutwin, C., Nevill-Manning, C.G.K.: Practical Automatic Keyphrase Extraction, Working Paper 00/5, Department of Computer Science, The University of Waikato, New Zealand (2000)
12. Witten, I.H.: Browsing around a digital library. In: Proceedings of Australasian Computer Science Conference, Auckland, New Zealand, pp. 1–14 (1999)
13. Turney, P.D.: Coherent keyphrase extraction via Web mining. In: Proceedings of IJCAI'03, pp. 434–439 (2003)
14. Turney, P.D.: Learning algorithms for keyphrase extraction. Information Retrieval 2(4), 303–336 (2000)
15. Desilets, A., de Brujin, B., Martin, J.: Extracting keyphrases from spoken audio documents, SIGIR Workshop on Information Retrieval Techniques for Speech Applications, pp. 36–50 (2001)
16. Inkpen, D., Désilets, A.: Extracting Semantically-Coherent Keyphrases from Speech. Journal of the Canadian Acoustics Association 32(3), 130–131 (2004)
17. Bain, K., Faisman, A.: Network ViaScribe Project. Liberated Learning Consortium (November 28, 2003)

Critical Success Factors for Automatic Speech Recognition in the Classroom

Steve Bennett[1], Jill Hewitt[1], Barry Mellor[2], and Caroline Lyon[1]

[1] Department of Computer Science, University of Hertfordshire, College Lane, Hatfield Al10 9AB, United Kingdom
S.J.Bennett@herts.ac.uk,
mail@SpeechSoft.co.uk.,J.A.Hewitt@herts.ac.uk,
C.M.Lyon@herts.ac.uk
[2] SpeechSoft Limited,The Carriers,Green End, Sandon, Buntingford, Herts, SG9 0RQ

Abstract. This study looked at continuous automated speech recognition (ASR) to an audience in a university lecture theatre and ran an evaluation based on a previous experiment by Ryba, McIvor, Shakir and Paez, which found that non-native speakers of English were much more favorable towards the use of ASR in class than native speakers. Our evaluation was done on a class of 29 students composed entirely of non-native speakers/ A strong indicator of the level of engagement with the technology was the linguistic ability of the user – the weaker the student's English, the more he or she tended to look at the textual output, the greater distraction experienced through poor recognition and also the greater impatience felt with slow recognition. There also seemed to be cultural differences – the Chinese students appeared to look at the textual output much more than Indian students. We conclude that the 2 axes around which successful classroom speech recognition occurs are those of accuracy and unobtrusiveness. The more accurate and unobtrusive the technology, the more successful will be the automatically transcribed lecture.

Keywords: Automatic Speech Recognition, multimodality, accuracy rate.

1 Introduction

One of the most fascinating recent studies done of speech recognition in class was Ryba,McIvor,Shakir and Paez's Liberated *Learning: Analysis of University Students' Perceptions and Experiences with Continuous Automated Speech Recognition*. [1] Using the Liberated Learning Consortium's tools, such as IBM ViaScribe™, they ran ASR over 3 lectures at Massey University NZ, and then an evaluation which contrasted specifically the opinions of native speakers as against non-native speakers of English, and discovered there a significant disparity between the favorability toward the technology of the two groups. This study not only evaluated how the students reacted to live ASR, but also what they thought of the subsequently edited captioned recordings of the event which were later made available to the students.

The Massey study was a very rigorous experiment involving taking the views of 160 students who had attended one or more of a series of three lectures. 81 of the

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 224–233, 2007.
© Springer-Verlag Berlin Heidelberg 2007

students were regarded as belonging to the L2 (English as a second language) group, and 79 to the native-English-speaking group (known as L1). Owing to the fact that it covered 3 separate lecturing events, we can be highly confident of the robustness of the findings.

The experiment described here was on a smaller scale but was deliberately analogous to the study of Ryba et al. We sought both to pose many of the same questions as asked in the Massey questionnaire, but also a set of other questions arising from the free-text feedback given by the students in that study.

Our study went to a class of 29 students of whom 17 were from India, 8 from China, 2 from Nigeria and 1 from Pakistan and 1 from Saudi Arabia. These numbers were gathered from self-declaration on their questionnaire. Regarding their proficiency at English language, 2 declared themselves to be lower intermediate, 11 described themselves to be intermediate, 11 described themselves as being upper-intermediate, and 3 described their level as advanced, one student did not respond to this question.

The lecture lasted 1 hour in which Dragon NaturallySpeaking version 8 speech recognition software was used, and at the end of it, a questionnaire was handed out containing 12 Likert scale questions asking their opinions on the use of ASR plus 2 more asking the students to describe their level of English and their country of origin.

Below is a comparison between the setup for the lecture series in the Massey study compared with our own:

Ryba,McIvor,Shakir & Perez	*Bennett,Hewitt,Mellor,Lyon*
1. Laptop computer	(THE SAME) In our case a Sony VAIO … CPU 1.7Mhz RAM 1MB
2. wireless microphone set	(THE SAME) in our case a Plantronics Bluetooth wireless microphone with inbuilt sound card (CS60-USB)
3.Viavoice 10™ – local voice profile	Dragon NaturallySpeaking 8 local voice profile
4.Viascribe™ display interface for automatically transcribing speech into text.	SpeechSoft SpeakView display interface software integrated with NaturallySpeaking
5. Text output to single in-class display via data projector	Slides were projected on one screen while text was projected on another
6. File storage on local hard disk	(THE SAME)
7. File transfer and editing	Done in Smirk (see text) – but N/A for this study
8. Lecture files uploaded to internal network	Done in Smirk – but N/A for this study

In our study there were were two screens being projected: one containing the recognized speech, the other the projected lecture. From which we can infer, it seems that one projector was used for both slides and text output (this is the implication of their *Figure 1* though it is not explicitly stated). In our case we used the software Smirk for the lecture rather than PowerPoint – owing to the fact that it records the sequence of slides together with the audio narration as well as pen movements written over the slide during the lecture. Smirk is a lecture capturing tool developed at the University of Hertfordshire which turns the audio recorded lecture into a sequence of jpg and mp3 files to be played back and streamed over the internet. [2]

Our trial took place in a room which had two data projectors built in – therefore the setup was presumably less onerous than that experienced by Ryba et al where the distracting quality of the setup was mentioned in the conclusions. In our case, the lecturer arrived with a laptop dedicated for speech recognition – and used the Plantronics Bluetooth microphone with it. This in earlier trials conducted by us seemed to give very good recognition, even with minimal training. In this case, the lecturer had simply run through the standard Dragon training procedure lasting no more than 30 minutes. No explicit prior rehearsal of the lecture had been attempted. In the lecture theatre itself, all that was done was to connect the laptop to the second projector, plug in the USB microphone, (checking the battery had been charged!), load (but not directly use) Dragon NaturallySpeaking, and then run SpeakView - our own text acquisition and presentation system which integrates with NaturallySpeaking.

SpeakView is an application which optionally uses PowerPoint. It has a resizable caption window which may occupy the whole display or in conjunction with PowerPoint become a caption window at the bottom of the screen which displays the recognized text. It also adds any user specified punctuation, such as full stops and new lines according to speech delays and has advanced text presentation options such as speed, color, scrolling styles and word/line spacing.[3] SpeakView can also broadcast to mobile phones and a WiFi network. In the experiment attempted here, it was put over a blank PowerPoint presentation on the one screen, while the original lecture text was projected from the other dedicated computer in the room onto a different screen.

Since our study was on a much smaller scale with fewer resources and personnel, allowing much less follow up, we did a less sophisticated questionnaire but one we believe addressed all the major points brought up in the Massey study. Here side by side are the directly related questions:

Ryba,McIvor,Shakir & Perez	*Bennett,Hewitt,Mellor,Lyon*
how much did you use the speech-text display *Not at all| Occasionally|Sometimes| Frequently| Nearly always*	I looked at the text display a lot during the lecture *Strongly disagree|Disagree|Neither agree nor disagree|Agree|Strongly agree*
The display helped me to understand the lecture Strongly Disagree|Disagree| Agree|Strongly Agree	The text display helped me to understand the lecture (options as above)
The display helped me to take notes (options as above)	The text display helped me to take notes (options as above)
I think most students can benefit from the Liberation Learning Project (options as above)	I think most students can benefit from the use of Speech Recognition in class (options as above)

In addition to these we also asked a number of questions arising from the free-text feedback in the Massey study:

Ryba,McIvor,Shakir & Perez			*Bennett,Hewitt,Mellor,Lyon*
Theme	Statements	+/-	Question (Likert Scale)
Visual	*See the words easily*	+	I was in a position in the class where I could easily see the displayed text

Distraction	*It was hard to read and served as more of a distraction*	-	The appearance of text on one of the screens all the time was distracting
Not accurate	*It would be useful if it was more accurate*	-	The speech recognition was accurate
Speed	*The speed of the system*	-	The time taken for the system to recognize the lecturer's words was too long
Colour	*Background colour should be darker (darker than lecture slides when in class) found it distracting in class*	-	The colour of the text on the screen was pleasant

The only questions we asked which did not have any real antecedence in the other study were

1. *The text display helped me translate the lecturer's speech* (Likert)
2. Level of English (Beginner, Lower Intermediate, Intermediate, Upper Intermediate, Advanced)
3. Country of origin

2 The Lecture Content

The lecture was an introduction to human factors in computing, specifically covering the ideas of Donald Norman and lasted 1 hour. There were 10 slides, and 5 occasions upon which the lecturer broke out of the slide system to draw on the tablet display of the dedicated computer. A number of the illustrations of themes contained well worn vignettes that he had delivered before while others were improvised at the time. The only real difference between this and the lecture he might have given without speech recognition was a greater deliberation in the words and the occasional accentuation of pauses. At the end of the lecture, the questionnaire was handed out at the end of the lecture to all the students seated in the class.

3 Evaluation of Results

In the section below, we will compare the results from the Massey study with our own, in the analogous categories. In the Massey study there was a great difference in results between the native speaking and the non native speaking subjects – and since all the members of the class in our experiment were non-native speaking – we therefore have decided the best comparison for our data would be with the *non-native English speaking segment* of the Massey study's student population. There were a total of 81 students belonging to the L2 non-native-English-speaking group, compared with our own 29.

Opinions regarding use of the speech text display:

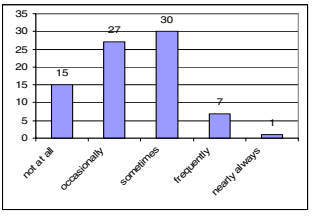


Fig. 1. Massey students (the L2 – non-native-English- speaking group) : ***how much did you use the speech-text display*** (from 80 answers) [in Ryba et al Table 1]

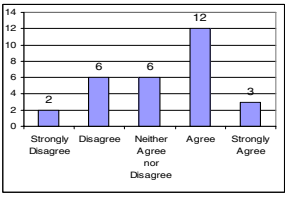


Fig. 2. Hertfordshire students: ***I looked at the text display a lot during the lecture*** (from 29 answers)

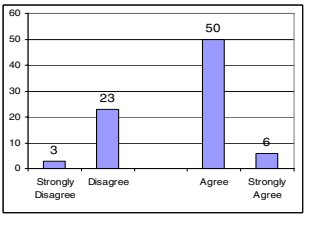


Fig. 3. Massey students (the L2 – non-native-English- speaking group): ***The display helped me understand the lecture*** (from 82 answers) [in Ryba et al Table 4]

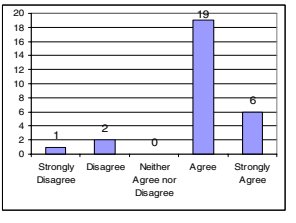


Fig. 4. Hertfordshire students: ***The text display helped me to understand the lecture*** (from 28 answers)

Opinions regarding note-taking and Speech Recognition generally

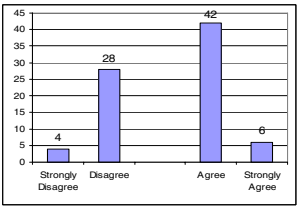


Fig. 5. Massey students (the L2 – non-native-English- speaking group): ***The display helped me to take notes*** (from 80 answers) [in Ryba et al Table 4]

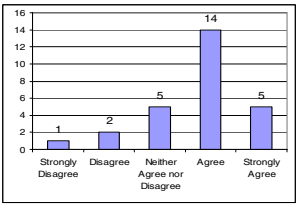


Fig. 6. Hertfordshire students: ***The text display helped me take notes*** (from 27 answers)

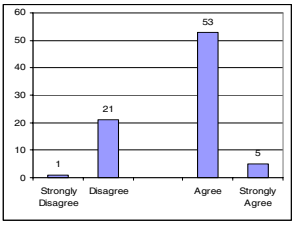


Fig. 7. Massey students (the L2 – non-native-English- speaking group): ***I think most students can benefit from the Liberation Learning Project*** (from 80 answers)

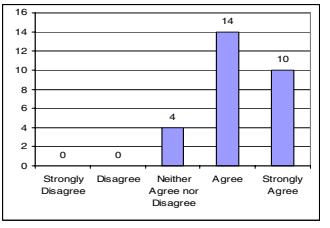


Fig. 8. Hertfordshire students: ***I think most students can benefit from the use of Speech Recognition in class*** (from 28 answers)

The greater favorability made toward our experiment (though over much smaller sample) is most likely caused by slightly better recognition and having two screens in the room and an easier setup. Caution must be expressed over these hypotheses however, because running a study over a much larger group makes it likely that:

1. Any *halo effect* is likely to be more pronounced on the smaller cohort since the relationship between the lecturer and the cohort is likely to be experienced as more immediate
2. The style of lecture is going to be different (less personal and more anonymous). One might also suspect that visibility of screens would be problematic for some of the attendees – however the treatment of this in the Massey study is mixed. While 17 students commented on the easy visuals, 31 commented about screens being hard to read and causing a distraction.
3. The smaller size of the Hertfordshire sample means less trustworthy results.

Similarly, the conclusions of the Massey study are likely to be more robust than ours, because over 3 lectures, one is more likely to obtain an average regarding lecturer performance, whereas in our own case, the positivity of some of the finding might only be a reflection of a lecture going particularly well.

Questions Arising from Free-Text Feedback in the Massey study

Regarding the potentially distracting quality of having the lecturers words outputted on screen, and levels of accuracy, the answers we received were:

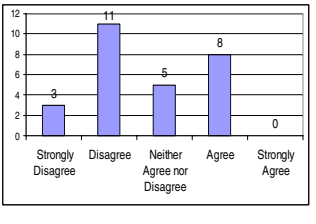


Fig. 9. Hertfordshire students: ***The appearance of text on one of the screens all the time was distracting*** (from 27 answers)

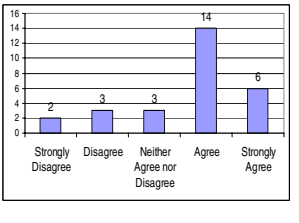


Fig. 10. Hertfordshire students: ***The speech recognition was accurate*** (from 28 answers)

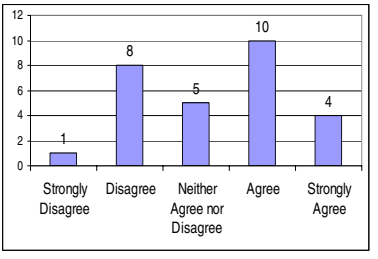


Fig. 11. Hertfordshire students: ***The time taken for the system to recognize the lecturer's words was too long*** (from 28 answers)

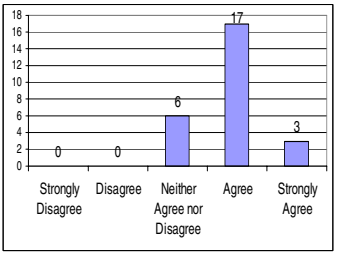


Fig. 12. Hertfordshire students: ***The color of the text on the screen was pleasant*** (from 26 answers)

4 Dependent Variables

In our analysis, there do seem to be some dependencies between the variables in this survey which we can look at through some cross-tabbing. Perhaps the most immediately striking is the difference between Indian and Chinese students. In this survey, there were 8 questionnaires received from Chinese students and 17 from Indian students. When their results are compared we find some interesting disparities:

Table 1. Difference between Indian and Chinese students

Statement	ALL	Indian	Chinese	Difference
I looked at the text display a lot during the lecture	3.28	2.88	4.13	1.25
The lecturer did not change his style to make use of the speech recognition	3	2.64	3.75	1.11
I would describe my level of English as (*Beginner, Lower-Intermediate, Intermediate, Upper Intermediate, Advanced*)	3.57	3.76	2.71	1.05

In this survey, there would seem to be a dependency between the variables relating to self estimation of language ability and country of origin. The Chinese students estimated their level to be considerably lower than that of the Indians. Perhaps this would explain why they seemed to look at the display much more than the Indian students. Equally there seems to be a big difference in the way they believe the lecturer changed his style to make use of the technology – the Indians believing he did change his style, the Chinese not. It is interesting to point out however, that the difference in attention paid to the text display did not show the same degree of difference when this measure was cross tabbed with general language ability (between the weaker English speakers and the stronger ones). Therefore, it does seem likely

that there are some cultural differences at work here as well as ones based on linguistic ability.

If do indeed try to look at the influence of language ability more exactly, and break up the cohort into two groups: those who consider themselves intermediate or lower in their language abilities (langA), and those who consider themselves upper-intermediate or advanced (langB), we find another interesting difference:

Table 2. Comparison of Students with Different Self-Declared Language Abilities

Statement	ALL	LangA	LangB	Difference
The appearance of text on one of the screens all the time was distracting	2.67	3.25	2.14	1.11
The lecturer did not change his style to make use of the speech recognition	3	3.54	2.33	1.21

(to précis: those with lower language ability *agreed the most* that the appearance of text *was* distracting, and *agreed the most* that the lecturer *had not changed his style*).

Therefore, it is those with lower language skills who found the text to be the most distracting – perhaps because they were depending on getting accurate recognition much more than those who did not require it. Similarly those with the lower language abilities most believed that the lecturer did not change his style to suit the technology. Perhaps this can be explained by the fact that since this group are struggling to truly understand the content, and therefore are less sensitive to questions of form and style than those who do not have any such difficulty with language.

5 General Observations

Once we factor out questions of accuracy, it seems the major variable controlling the perceived value of automatic speech recognition in class is the language ability of the individual student – with some element of cultural difference between Indian and Chinese students also visible. However, in general terms we can say, the weaker the student's grasp of English, the more attention they will pay to the textual output display, the greater distraction they will experience through the misrecognition of words, the greater impatience they will feel if there is a long latency between utterance and its recognition, and also the greater tolerance or even non-cognizance of variation in lecturer style owing to the presence of the technology. That is to say the students who can most benefit from this technology, will be the most demanding of it, and the most critical of its shortcomings. Conversely those who were more confident in their English language ability had very little issue with the intrusive nature of the speech recognition – in most cases it seems, they simply ignored the display.

This is quite different from the results of the Massey study where considerable negativity seems to have been registered by the native speakers. A potential explanation might be that, in our case, looking from the visual perspective of the students, the slides of the presentation as well as the lecturer were on the left of the room and the text display was on the screen to the right of the room.

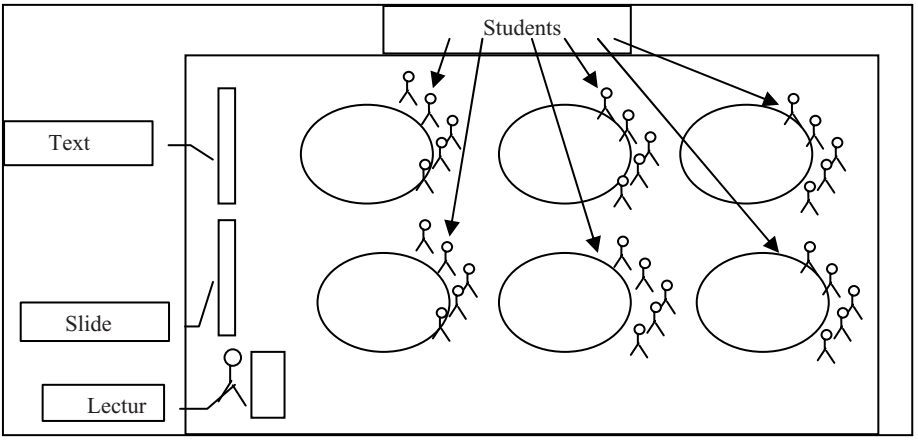


Fig. 13. Layout of the teaching accommodation

This would mean that those students who did not wish to look at the text display and be distracted by it, could just focus on the lecturer instead.

6 Conclusions vis-à-vis the Massey Study

The authors make 8 recommendations. I will go through the first four which relate to the practicalities of speech recognition – since the final 4 have a pedagogic flavor which is beyond the scope of this study.

1. Improved speech recognition – While Ryba et al advance more training as a way to achieve this, we believe the quickest way is to get a good microphone. Our own success with recognition improved enormously after using the Plantronics CS60-USB microphone. This improved our recognition rates enormously. The other intervention involves playing with the settings – both Dragon and ViaVoice have controllers which specify the tradeoff between accuracy and speed. In our experience, slowness seems to be forgiven far more easily than mis-recognition. An unavoidable adjustment does need to be made in the lecturer's style, particularly if there is some latency in the recognition. Good recognition requires a more deliberate delivery and longer pauses. The challenge is for the lecturer to integrate speaking like this with a discursive mode which still remains undeniably theirs.
2. *Improved setup procedure that is less distracting* – this may have been caused by the large cohort size in the Massey study. Our study was facilitated by the fact we used a room which had two data projectors built in. We also began at 9.00 and were able to occupy the room before the students arrived. Certainly it seems that having two built-in projection screens in one of the superior teaching rooms of the university was a major help in the success of our experiment.
3. *More interactive features to improve communication with* students – Ryba et al talk of potential learning scenarios, for instance collaborating as a class to list points for

discussion which then appear on the screen. We think this is an interesting idea, but we again wish to make sure that academics do not have to adjust their practices unnecessarily in order to make use of the technology

4. *Seating arrangements so that there is a specific area where people who wish to see the speech-text screen can chose to sit in that area.* As we mentioned above, the layout of our room meant that for most of the students, the recognized text was not foregrounded. Students could divert their attention to the text display as and when necessary.

This much smaller scale experiment seems bear out and maybe put in sharper relief the majority of the findings of the Massey study. There are however a number of areas where further research is necessary.

Firstly, studies need to be made which correlate these measures of student satisfaction against true measures of recognition accuracy. In our study, unfortunately the software crashed at the moment of saving the transcription of the lecture and so we could only compare student opinions against their *felt* perception of the recognition accuracy, not against an absolute measure of that accuracy itself. Secondly a study such as this needs to be carried out on a larger cohort and over a longer period of time, as was the case for the Massey study, to check whether the more favorable results we obtained as compared to their study might merely be the reflection of more intimate surroundings and potentially superior teaching accommodation and some luck *on the day*.

However, we can conclude that the critical success factors for successful speech recognition in class are 1) level of accuracy of transcription 2) unobtrusiveness of the technology, both at the level of setup and at the level of visualization. If these two factors are attended to, ASR is a technology which can seriously improve the learning experiences of students in the lecture theatre.

References

1. Ryba, K., McIvor, T., Shakir, M., Paez, D.: Liberated Learning: Analysis of Students Perceptions and Experiences with Continuous Automated Speech Recognition. The Electronic Journal of Instructional Science and Technology, vol. 9(1) (2006)
2. Bennett, S., Hewitt, J., Kraithman, D., Britton, C.: Making Chalk and Talk Accessible. In: Proceedings of the 2003 conference on Universal usability (Vancouver, Canada), ACM Press, New York (2003)
3. Hewitt, J., Lyon, C., Britton, M.B.: SpeakView: Live Captioning of Lectures. In: Proc. HCI International, vol. 8 (2005)

Developing and Evaluating a TriAccess Reading System

Ming-Chung Chen[1], Chien-Chuan Cko[2], Lih-Yuan Chen[1], and Chun-Han Chiang[1]

[1] National Chiayi University, Department of Special Education
[2] National Chiayi University, Department of Computer Science and Information Engineering
300 University Rd., Chiayi, Taiwan
{mtchen,cko,lihyuanchen,s0940606}@mail.ncyu.edu.tw

Abstract. Learners with disabilities experience many difficulties when reading. This study aimed to develop a TriAccess computer system to assist learners with disabilities to read effectively. TriAccess provide physical, sensory, and cognitive supports to learners. The system also offers professionals a convenient and web-based interface on material development and system management. In addition, the study conducted a usability evaluation on the TriAccess. Thirty participants involved in the evaluation process. The results have shown that the TriAccess system is valuable not only to learners, but also to instructors, and material developers.

Keywords: reading difficulties, universal design, multiple representations.

1 Introduction

Students learn reading skills since they were young. Reading is not only an essential skill but an important activity for learning. They learn new knowledge, events or concept through reading the textbooks, reference books and related websites. Effective reading, however, simultaneously requires a learner's competences on physical manipulation, sensory perception and cognitive processing. Obviously, it is impossible or difficult for learners with disabilities to read effectively without any help, because of their limitations or disabilities, such as dyslexia, tremble, visual impairment, visual perception difficulties, or palsy.

For years, professionals have been trying hard to find solutions to assist those who experience difficulties in reading. In addition to remedying the impairment or improving the skill, providing assistive technology device is another solution to comprehend text by mean of bypassing people's disabilities or augmenting their residual capabilities [1]. For instance, reading aloud program can help poor readers to learn electronic text by listening; screen magnifier in the computer system can assist people with low vision to be able to read electronic text by enlarge the characters.

Generally speaking, assistive technology devices provide solutions to solve the problem of "access to material," instead of the problem of "access to learning" [2]. For students with cognitive limitations, however, they can not read effectively only by accessing to materials. In addition, they also need cognitive supports to learn through reading. Therefore, solutions with various reading supports are in need for students with different types of reading difficulties.

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 234–241, 2007.
© Springer-Verlag Berlin Heidelberg 2007

Fortunately, Universal design has been introduced for years in the field of architecture, and it has already applied to education recently [2],[3],[4],[5]. Universal Design for Learning (UDL), advocated by Center for Applied Special Technology (CAST), is one of the applications in education [2]. UDL emphasizes that all learners with or without disabilities should reach the same learning objects by methods which are appropriate for them. This idea is especially important for learner who receive challenging curriculum in inclusive education environment. This approach also expands our vision on how to help students with reading difficulties.

The authors adopted one of the principles of UDL, multiple representations, to develop a supportive reading environment, TriAccess reading system, which could provide physical access, sensory access, and cognitive access for readers. Furthermore, researchers also evaluate the system usability on TraiAccess.

2 System Developing

2.1 Features Analysis

After reviewing previous related studies [2],[6],[7], the researchers in this study proposed 16 possible solutions for reading supports. As shown in Table 1, there are several possible difficulties resulting from the accessibilities on physical, sensory, and cognitive problems. Accordingly, the researchers proposed 16 useful strategies to dealing with these difficulties. A Delphi survey was conducted to validate these strategies. Seventeen professionals, from curriculum design, special education, and instruction fields, were invited to participate in the Delphi survey. They were asked to rate the importance of the 16 potential strategies in the list. After two rounds of Delphi survey, there were 14 items regarded as 'important' or 'very important' strategies. Therefore, these 14 strategies were then adopted in the TriAccess.

2.2 System Creating

TriAccess is a web-based reading system. It allows a convenient way for system manager, teachers and learners to finish their tasks through the Internet. TriAccess was developed in Windows Server 2003. ASP.NET and JAVA Script were used to develop the user interfaces and the management system. SQL was adopted to store the curriculum material and user profiles, including the specific configurations for individual learner's reading environment.

This system provides multimedia to support users' reading. In order to integrate speech synthesis systems into TriAccess, authors provided Microsoft Agent, Visual Basic ActiveX, and Character MP3 (a popular speech synthesized program in Taiwan) as options for learners to select. However, learners need to install the program on their computer in advance. TriAccess could read aloud whatever the text marked by reader without operating extra speech synthesized program. Besides, TriAccess adopted Stream Media technique to display the video effectively and Cascading Style Sheets (CSS) technique was conducted to provide flexible appearance of text.

Table 1. The usefull strategies for sloving the difficulties of accessibilities

accessibilities	Possible difficulties	Useful strategies
Physical access	● Postural position ● Motor control of upper limps ■ Gross motor ■ Fine motor ● Lacking upper limps	● Flexible interface ■ Location of icons ■ Sizes of icons ■ Interval between icons
Sensory access	● Low vision ● Discriminating front from background poorly ● Blind	● Magnifying size of text ● Changing font of text* ● Contrasting the color of text with background ● Increasing spacing of words and lines ● Setting a sentence per line ● Putting numbers at the beginning of each line* ● Reading aloud on demand
Cognitive access	● Attention & perception ● Word- recognition ● Comprehension ■ Syntax ■ Concept of keyword ■ Background knowledge ■ Working memory	● Marking keywords or key concepts ● Explaining Keywords in various representations ■ Text, voice, picture, and video ● Providing background knowledge ● Assisting holistic understanding ■ summary ■ concept map

Note: * not regarded as important strategies by the professionals.

2.3 System Overview

TriAccess includes 'system manger' and three kinds of user interfaces for material developers, instructors, and learners. After users enter their login identification and password, the system brings up the authorized interface accordingly. Generally, an assigned material developer uploads the text and required cognitive supports on material developing interface first. The system then organizes these supports automatically and stores them in the database. Secondly, the authorized instructors set up an individualized reading environment for their students (learners). The last, the learners can access the particular subjects for their reading and with individualized supports. The major functions of the interface for each kind of users are as following:

Interface for Material Developer. First of all, material developers need to find reading materials and digitalized the materials in TriAccess. As shown in Figure 1, TriAccess provides friendly interface for material developers to prepare the reading materials. The digitalized materials include copying and pasting the reading texts, and the summary, the concept map, and the background knowledge of the texts. Within the reading texts, material developers also mark keywords and explain these keywords in texts, voices, pictures and video clips. These are all cognitive supports mentioned in Table 1. The developers click on 'upload' button and finish their work.

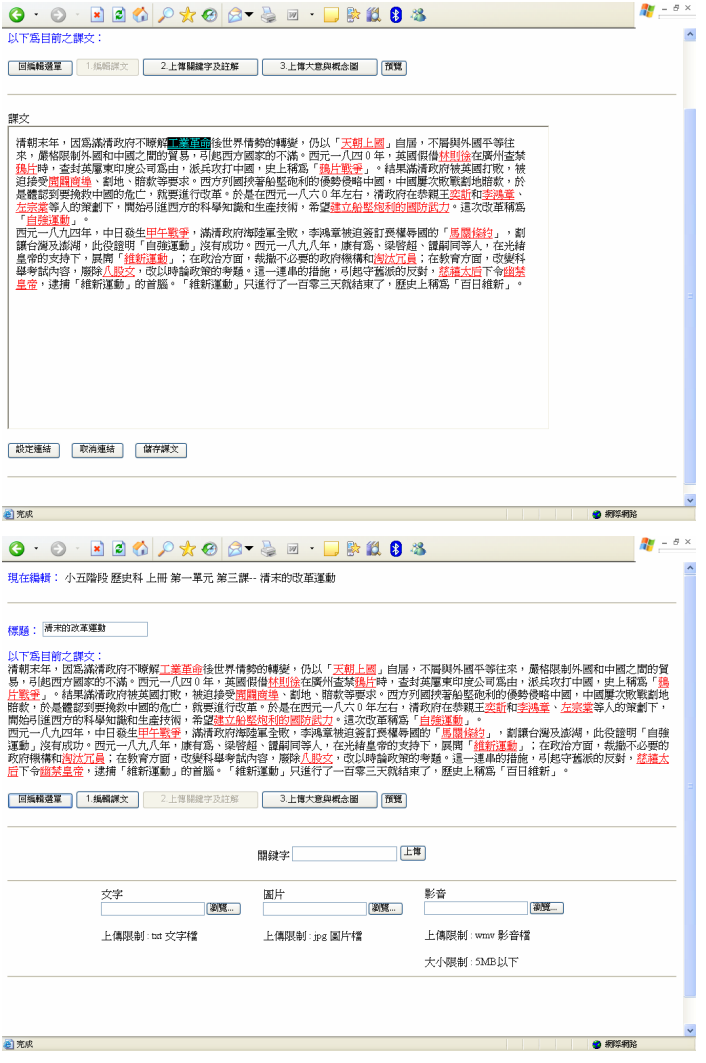


Fig. 1. The screen shots of the interface for a material developer to edit a lesson. The upper one is used to mark key concepts; the lower one is used to upload cognitive supports for specific key concept.

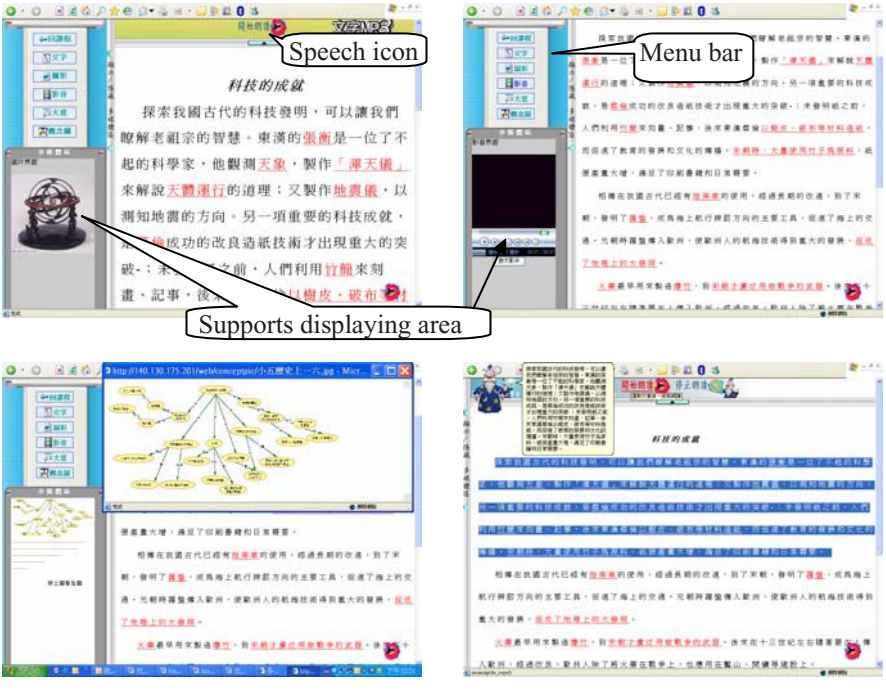


Fig. 2. Examples of Learner's interface. The article of these examples is identical but displayed in different conditions and has shown in four frames. The upper left frame reveals picture support of key concept and larger spacing and font size; the upper right one demonstrates video support of key concepts; the text of lower left one displays a concept map of the article; in the lower right one, the menu bar and cognitive support displaying area are disabled, and Micro Soft Agent-Merlin is activating.

Interface for Instructor. Providing readers a challenging but supportive reading environment can stimulate them to read effectively and to improve their reading skill as well [2]. Therefore, it is important to select individualized supports for readers. TriAccess offers the interface for instructor to do this task.

Instructors set up their students' profiles first. The profile includes each student's name, grade, and disability etc. The instructors then assign each student the reading materials and the needed physical, sensory, and cognitive supports (as mentioned in Table 1) based on their students' reading abilities and preferences.

Interface for Learner. The purpose of TriAccess system is to provide physical, sensory and cognitive access supports to the readers. After the material developers prepare the digitalized reading materials and the instructors assign their students' individualized supports, the students can login learner interface, click on any of the assigned materials, and then TriAccess brings up the reading text with all the supports they need. Figure 2 shows how the learner's interface looks like.

3 Usability Evaluating

3.1 Method

Thirty participants involved in the system usability evaluation. Five of them were faculties in universities, four were experienced school teachers, 13 were practice special education teachers, and eight were pupils with special needs.

The participants were asked to test TriAccess system through three different interfaces, including material developer, instructor, and learner. The professionals, practice teachers and students operated the interface for learners only, while the experienced teachers evaluated the interfaces for developers, instructors and learners. The faculties and teachers filled out a five-point rating scale questionnaire regarding the usability of the TriAccess. The usability includes satisfaction, remembrance, efficiency, recognizability, and learnability. Teachers also read a student version questionnaire (with satisfaction, interesting, remembrance, and benefit dimensions) to their participated students and the students answered orally.

3.2 Results

Participants need to rank each description from very agreement to very disagreement in the five-point rating scale questionnaire. A five-point means "very agreement" where as one-point means "very disagreement. The mean of each dimension was used to represent the usability degree. Table 2 only presents the results of usability evaluation by 22 participated professionals, not including the data from participated students. Students' data will be reported in the 3.2.3 section.

Interface for Material Developer. The experienced school teachers evaluated the interface for material developer. As the results in table 2 indicated, the score of rating was higher than 3, regarded as indicator of agreement. Teachers tended to agree the interface designed for material developer was usable. In addition, the teachers provided some suggestions for improving the functions of the interface, such as adding the function of preview to let the developer check the lesson before publishing it.

Interface for Instructor. The experienced school teachers also evaluated the interface for instructor. As the results in table 2 shown, the score of rating was higher than 3, too. Teacher tended to agree the interface designed for instructor was usable. Besides, they suggested simplifying the interface.

Interface for Learner. All the 30 participants evaluated the interface for learner. As the results released in table 2, the score of rating, on average, was higher than 3. The professionals and the teachers tended to agree that the usability of TriAccess system is good. The professionals also have a higher score than the teachers in all five dimensions. Besides, the eight participated students were asked to evaluate the interface for learner. They expressed their agreement, too (interesting=3.8, remembrance = 3.5, satisfaction = 3.9, benefit = 3.5). They thought that TriAccess was interesting and useful for them.

4 Discussions and Conclusions

This study aimed to develop a multi-support reading system based on the solutions provided by related studies. TriAccess not only provides individual learner appropriate physical, sensory, and/or cognitive supports depended on her/his needs, but provides text material developers and instructors a simple and convenient interface to manage teaching material.

Table 2. The score of usability evalation of three interfaces

Interface	Evaluator	Satis.	Remem.	Effic.	Recog.	Learn.
Developer	Teacher	3.6	3.8	4.1	3.9	3.6
Instructor	Teacher	3.8	3.8	3.9	3.6	3.5
Learner	Professional	4.1	4.4	4.5	4.0	4.2
	Teacher	3.5	3.8	3.8	3.8	3.4
	Practice Teacher	3.8	3.8	4.1	3.7	3.8

In sum, TriAccess offers at least three advantages for users. First of all, it provides different reading access solutions for learners. This function is helpful to students with disabilities who participate in general education curriculum and have difficulties when access to printed reading materials. Secondly, teachers may develop their teaching material easily with many supports which their students need, especially for cognitive supports. Finally, TriAccess also provides teacher a simple interface to set individualized reading environment for each student easily. Then, the student will be able to read through their own way.

Besides, the results of usability evaluation has shown that the interface of material developers, instructors, and learners were not difficult to manipulate. The results of usability evaluation also indicated that efficiency gained highest score. It expressed that the evaluators agreed that TriAccess could increase the production for material developers, instructors and learners. Therefore, the TriAccess system responded the major purpose of this study. However, the simplification of the interface was not enough. Therefore, the researchers had adopted some suggestions to revise the interface, such as adding the function of preview for material developer, simplifying the access options setting interface for instructors.

Although experience school teachers simulated the process of editing a lesson when they evaluated the usability of the interface for material developer in this study, they did not really develop a lesson of text with cognitive supports and upload the text and related elements. In the future, the researchers have planed to host a workshop for teachers to learn how to develop reading materials in TriAccess, and then ask these participated teachers to do the evaluation.

In addition, it is also in need to conduct an experiment to explore the effectiveness of using TriAccess system for students with disabilities while they actually participate in reading activities in the classroom. The researchers also have interests to explore the benefit of specific support for students with specific disability. The research-based evidence will assist instructors to select the suitable supports for their students.

Acknowledgments. The authors would like to thank the National Science Council of the Republic of China for financially supporting this research under Contract No. 95-2614-H-415-001-F20.

References

1. Lewis, R.B.: Special Education Technology Classroom Applications. Brooks Cole, Pacific Grove CA (1993)
2. Rose, D.H., Meyer, A.: Teaching Every Student in the Digital Age Universal Design for Learning. Association for Supervision and Curriculum Development, Alexandria VA (2002)
3. Orkwis, R.: Universal Designed Instruction. ERIC/OSEP Digest(E641). ERIC Cleaninghouse on Disabilities and Gifted Education (2003)
4. Scott, S.S., McGuire, J.M., Shaw, S.F: Universal Design for Instruction: A New Paradigm for Adult Instruction in Postsecondary Education. Remedial and Special Education 24(6), 279–369 (2003)
5. Silver, P., Bourke, A., Strehorn, K.: Universal Instructional Design in Higher Education: An Approach for Inclusion. Excellent. Equity in Education 31, 47–51 (1998)
6. Chu, C.N., Li, T.Y., Chen, M.C.: The Design of an Adaptive Web Browser for Young Children with Reading Difficulties. In: Miesenberger, K., Klaus, J., Zagler, W. (eds.) ICCHP 2002. LNCS, vol. 2398, pp. 189–190. Springer, Heidelberg (2002)
7. LoPresti, E.F., Mihailidis, A., Kirsch, N.: Assistive Technology for Cognitive Rehabilitation. State of the Art. Neuropsychological Rehabilitation 14(1/2), 5–39 (2004)

CAT Motor: An Innovative System to Detect the Behavior of Human Computer Interaction for People with Upper Limb Impairment

Chien-Chuan Cko[1], Ming-Chung Chen[2], Ting-Fang Wu[3], Su-Ying Chen[2], and Chih-Ching Yeh[4]

[1] National Chiayi University, Department of Computer Science and Information Engineering
[2] National Chiayi University, Department of Special Education
300 University Rd., Chiayi, Taiwan
{cko,mtchen}@mail.ncyu.edu.tw,
mzqp@ice.ntnu.edu.tw
[3] National Taiwan Normal University, Graduate Institute of Rehabilitation Counsel
[4]National Taiwan Normal University, Department of Information & Computer Education
162, Sec. 1 Hoping East Rd., Taipei, Taiwan
tfwu@ntnu.edu.tw,
dexteryeh@mail.cy.edu.tw

Abstract. This paper described a system which could measure cursor movement and control site motion simultaneously. In addition to evaluate a client's pointing and selecting proficiency, CAT Motor system, developed on the basis of the previous CAT system, was also used to detect the motion of his control site. This feature could provide clinical professionals more detailed information when comparing the difference of performance among devices. The CAT Motor system consisted of four basic types of mouse operating skills, including clicking, cursor moving, cursor moving and clicking, and dragging. The CAT Motor recorded the location and the response of the pointing device and captured the picture of the control site synchronally. Then, the system visualized the results of cursor moving and control site motion on the screen. The CAT Motor system not only provided the professionals with an objective mean of evaluating specific difficulties of mouse operating difficulties, but also allowed them to evaluate the effort of the motor control training. Besides, this paper had some suggestions for the research in the future.

Keywords: computer interaction, motion detection, evaluation system.

1 Introduction

With the rapid growth of computer technology, the use of computers has been an essential part in our daily lives. For those with physical impairment, a computer is not only a novel product but a necessary tool for education, occupation, communication, and rehabilitation. In order to satisfy the needs, researchers have participated in

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 242–250, 2007.
© Springer-Verlag Berlin Heidelberg 2007

innovative human computer interface studies, including interacting technique and devices. Moreover, researchers explore the behavior of human computer interaction to find out the specific difficulties that people with physical impairment faced and determine the useful interventions for them. In the past behavior studies, researchers analyzed the performance of the mouse operating task whatever the parameters were adopted [1],[2],[3],[4],[5].

Fitts' Law was used to explore the behavior of computer interaction widely [6]. Throughput had been regarded as a new standard indicator to measure the effect of mouse recently (ISO 9241-9). However, these indicators could show the existence of the difference only, rather than the reason of the difference. Therefore, in addition to speed and accuracy, Mackenzie, Kauppinen, and Silfverberg, (2001) brought up seven new parameters as new measurement indicators, including target re-entry (TRE), task axis crossing (TAC), movement direction change (MDC), orthogonal direction change (ODC), movement variability (MV), movement error (ME), movement offset (MO) [3]. Keates, Hwang, Langdon, and Clarkson (2002) added missed click (MC) and ratio of path length to task axis length (PL/TA)[2]. They tried to explore the reason that result to the difference between classes of device and the difference within classes of devices. Nevertheless, they did not investigate the motor performance of the client's control site simultaneously, nor did the authors [7].

The authors developed a computerized assessment tool for mouse proficiency, named as CAT, to assist the clinical rehabilitators to evaluate and select suitable pointing devices for their clients [7]. Although CAT was useful for the rehabilitators, they found some situations that could not be detected by the CAT system. Take a client for instance, she completed a pointing task with a grouped switches mouse twice in different conditions. In the first condition, she reached her fist to the mouse quickly but aims to the specific switch unstably when the mouse was located near her side. In the second condition, she needed more time to raise her arm to reach her fist to the mouse but can press the target switch immediately. Although the reaction times recorded in the two conditions were identical, the processes were totally different. The difference between them would not be discriminated until the performances of the client's fist movement were analyzed.

The technique of motion detection has been implemented in human computer interaction, especial for alternative input technology [8]. It captures the movement of a specific control site with sensors or video cameras and they will be translated into the recognizable signals. Motion detection is widely used as the key technique for moving object extraction from image frames. It can be utilized to measure velocity or track the trajectory of moving object. The technique of movement tracking may be used to detect the motor movement of client's control site. It could provide rehabilitation professionals more detailed and useful information related to selecting more suitable interacting devices. Furthermore, a system which incorporates cursor and motion detection could let evaluators examine the improvement of a motor control therapy by comparing the motion trajectories. Therefore, this paper aimed to develop a synchro-assessment system of mouse operating proficiency which measures cursor and motion trajectory simultaneously.

2 System Development

2.1 Idea of System Design

Based on the aforementioned purpose, the authors tried to create a system named CAT Motor that could measure the movement of cursor on the screen and the motion of control site that generates the cursor simultaneously. As the framework shown in figure one, CAT Motor consists of four major modules and a database. The CAT Motor was designed on the basis of the clinical approach. It adopted client-based data management mode to create database.

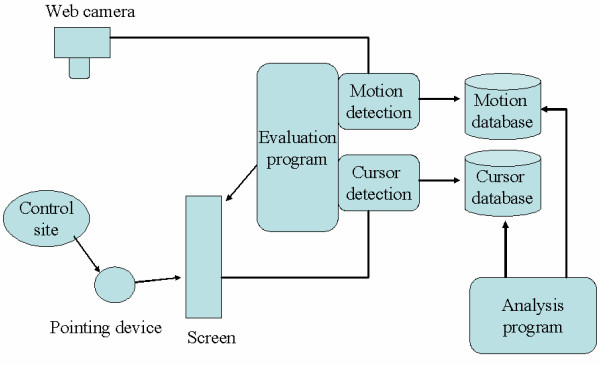


Fig. 1. Framework of CAT Motor system. This shows the four major program modules in the system.

The client sat in front of the experiment desk. He used his control site to operate a pointing device. For example, he operated a trackball with his right fist. The evaluation tasks displayed on the screen once the evaluation started. The client was required to complete the tasks as correctly and quickly as he could. The motion detection program would request the camera to capture the picture and store them into the database. Meanwhile, the cursor detection program would record the cursor's trajectory and the response of selection. The motions of the control site and the cursor were recorded synchronally until the task was completed. The clinical professional could use the analysis program to review the results of cursor movement and the motion of the control site on the screen after the evaluation was finished.

2.2 Method of System Development

Based on the previous effort CAT system [7], authors adopted the kernel of CAT system as the basis to develop this CAT Motor. The essential tasks for computer interaction which were included in the sub-assessment systems of Basic Skills (CAT-BS) and Mouse Proficiency (CAT-MP) were used to develop the evaluation program. The evaluation program enables evaluators to setup the specific tasks for the

measurement. Moreover, when the measurement starts, the program synchronizes the motion detection program and the cursor detection program to record the data.

The cursor detection program detect the location(x-coordinate and y-coordinate) of the cursor every deci-second that could be used to review the cursor trajectory and to compute some essential parameters. For example, movement time (MT), throughput, target re-entry (TRE), movement variability (MV). In addition, the program records the response of the selection which could be used to compute the accuracy of each evaluation task. These data were stored in the database.

To capture the image, the web-cam had to be activated with motion detection program. Since there were some Active controls available, the authors selected VedioCap Live v2.0 Active Control to activate the web cam to capture the image of the client's control site. Although VedioCAp Live could capture video, the purpose of this system was to detect the trajectory of control site movement. Image could be analyzed more easily when calculating the position of the control site.

Therefore, VedioCap Live was integrated into the motion detection program to activate the web cam to capture the picture per deci-second until the client complete the evaluation. For reducing the loading of the processor, the size of the image was fixed to 320* 240. In addition, the detection program did not analyze the image. The data would be dealt with the analysis program.

Analysis program provided an interface to review the performance by generating data from the database. Besides displaying the result quantitatively, the program can synchronously display the trajectory of cursor movement and the control site as well as the pictures of the motion of the control site on the screen.

A circle sticker with specific color, light green was used as an example in this paper, was pasted on the client's control site which was used to access the device for representing the location of the control site. The central part of the sticker was used to represent the coordinate of the sticker. The average of the coordinate of all pixels in the sticker, (P_X, P_Y), was regarded as the central part of the sticker. The formula used to calculate the location of the circle sticker in every image is listed below:

$$P_X = (\Sigma P_{xi}) / n \quad (1)$$

P_X represents the position of the circle sticker in every image. P_{xi} represents the x coordinate of the every light green pixel in the image. "n" represents the amount of the light green pixels in the image.

$$P_Y = (\Sigma P_{yi}) / n \quad (2)$$

P_Y represents the Y coordinate of the circle sticker in every image. P_{yi} represents the Y coordinate of the every light green pixel in the image. "n" represents the amount of the light green pixels in the image.

The program can get the central coordinate of the circle sticker in every image and draw the trajectory of the control site's motion as well as the trajectory of the cursor movement simultaneously.

The authors adopted C++ Builder to develop the main four program modules of CAT Motor, and the database was created with Microsoft Access.

3 Feature of System

As the aforementioned framework, CAT Motor includes four main program modules. The major feature of each module will be described below:

3.1 Evaluation Program

The evaluation program compasses four tests: targeting, stationary clicking, targeting and clicking, and dragging. As the table one indicated, the evaluation tasks were originated from the CAT-BS and CAT-MP which were designed on the basis of the result of tasks analysis of essential mouse proficiency [7]. Within each test, the evaluator could flexibly select the target size, targeting distance, and targeting direction from the evaluation tasks setting menu. It allowed the professional to manipulate the index of difficulty (ID) for the evaluation task to compare the performance among different IDs. In addition, the evaluation tasks can be uniquely arranged depended on the client's proficiency or purpose of the assessment. Take the task "targeting and clicking" for instance, a target icon with a diameter of 1cm would be selected and compared with the performance of the different distances, 10 cm and 5 cm. The assessment could focus on evaluating the distance of cursor moving that a client can do. Conversely, the evaluator could fix the moving distance and change icon size if he wanted to explore the most appropriate size for a client to click the mouse button. Moreover, the program can synchronize the motion detection system and the cursor detection system by activating the start icon.

Table 1. Descriptions tests and their evaluation tasks

Tests	Tasks
■ targeting	Cursor moving
■ stationary clicking	L-single, L-Double, and R-single clicking
■ targeting and clicking	Cursor moving and L-single, L-Double
■ dragging	Moving icon to a target area

3.2 Motion Detection System

As the Fig. 1 shown, motion detection system includes a web camera and a motion detection program. The web camera captures the picture of a client's control site and translates the signal via a standard universal serial bus (USB) computer interface to the database. The program captured the image per deci-second and stored them into the database for analysis later.

In order to capture the whole picture of the control site, the camera was fixed above the evaluation desk, as the Fig. 2 indicated. Furthermore, professionals can modify the height of camera to explore a most suitable area and capture the client's motion if needed.

The motion detection program was executed in the background during the evaluation. The client who took the evaluation would not see the information so that he can pay his attention on the evaluation tasks.

Fig. 2. The situation of the evaluation. The height of the camera could be modified if needed.

3.3 Cursor Detection Program

In order to compute the speed, accuracy, trajectory, and other essential parameters mentioned by Mackenzie et al. [3], the cursor detection program was executed in the background during the evaluation. The client was not aware of the diction so that he could concentrate on his evaluation task. Once each evaluation task was started, the program detected the location of the cursor every deci-second until the task was completed. In addition, the response of the selecting would be recorded once a client pressed the button of selection. It was a signal for the CAT Motor system to terminate the program of evaluation, motor detection, and cursor detection simultaneously.

3.4 Analysis Program

Since both the cursor detection program and motion detection program recorded the data only during the evaluation process, the analysis program is responsible to the compute works. Rather analyzing a group subjects' average performance, the analysis

module is designed to overview the results of each client's performance in a single evaluation or the difference between two evaluations.

The analysis program adopted some essential parameters to serve as the indicators for comparing the performance between different tasks, pointing devices, and operating environments. The parameters comprise accuracy, speed, index of difficulty (ID), index of performance (IP), ration of actual moving path and distance of a task (RPT), numbers of submovements (SM), peak of velocity (PV), movement variability (MV), and numbers of target reentry (TRE).

Accuracy is the most important element when considering the performance. Since the analysis here is task-based, the result of the accuracy displays in "true" or "false".

Speed means movement time in some analyses. However, the speed is the task complete time (TT) spent for accomplishing a single test item correctly in this paper. As the equation 3 indicated, total time consists of three parts. Reaction time (RT) is measured from a target displayed on the screen for selection to the time the client activated the pointing device to move the cursor. In other words, the duration of the speed of the device would be transferred from zero to non-zero. The movement time (MT) is the duration the cursor spent for reaching the target. The correction time (CT) is measured from the movement of the cursor in the target to the time the client activated the selection button.

$$TT = RT + MT + CT \tag{3}$$

Index of difficulty (ID) has linear relationship with task complete time, according to Fitts' Law. The equation 4 describes the function of ID. It is calculated from A, the distance to the target, and W, the width of the target. The diameter of the target was used as the width of the target in this paper.

$$ID=\log_2(A/W+1) \tag{4}$$

In stead of exploring the effect of a pointing device, the CAT Motor purposed to analyze the result of the evaluation for each client. The author adopted index of performance (IP) instead of throughput (TP) to express the production [9]. As the equation 5 revealed, it represents the distance of the movement of the cursor per second.

$$IP = ID / MT \tag{5}$$

Ration of actual moving path and distance of a task (RPT) was used to explore the efficiency of the cursor movement. The analysis program calculates the distance of the movement of the cursor. Then the distance will be divided with the distance of the task. The optimum RPT is equal to 1, or less than 1, if the cursor just moves into the area of the target.

Numbers of submovements (SM) is computed within each trial. Different from the study of Hwang and his colleagues, they used three situations as considering the submovement [1], however, this paper adopted one of the criteria only. A submovement is counted when the speed of cursor motion changed from greater zero to zero within deci-second.

Based on Mackenzie and his colleagues' definitions [1], target re-entry (TRE) occurred when cursor enter the target, leaves and re-enter the target. Movement

variability (MV) represents the extent to which the sample cursor points lie in a straight line along an axis parallel to the task axis.

The result of those aforementioned parameters would display on the screen for each trail. CAT Motor system provides two options for the professionals to review the result. The first one is the cursor movement mode. It enables the professionals to compare the performance between two trails. As the Fig. 3 shown, in additional to displaying the quantitative results, it provides the visualized graph. The second mode is cursor and motor motion mode. The program will draw the trajectory of the cursor and the control dynamically, the images of the control sites as well. As the Fig. 4 indicated, the professionals could review the motion of cursor and control site simultaneously.

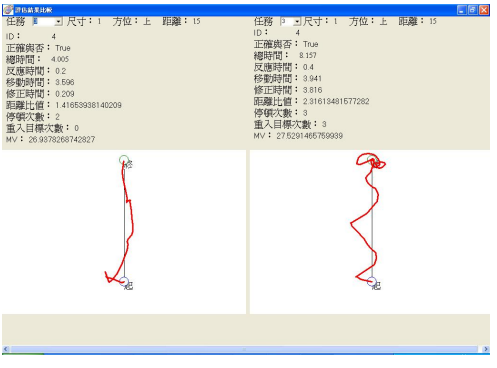


Fig. 3. The screen shot of the cursor movement mode

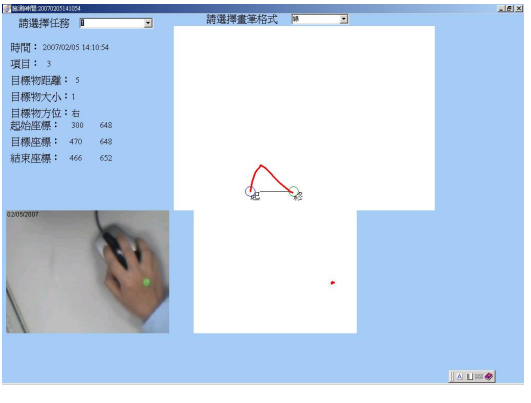


Fig. 4. The screen shot of the cursor and motor motion mode

4 Conclusions

This paper describes the major features of a computerized mouse proficiency assessment tool that combines traditional cursor diction and motion detection. The CAT Motor system was developed by embedding the motion detection function with

the kernel of the previous CAT system. Through the CAT Motor system, the clinical professionals could not only measure a client's computer interaction performance, but understand the motion of the control site. The result of the evaluation could assist the professional to select a suitable pointing device and the location to set the device as well.

Furthermore, the clinical professionals may need to discuss with other experts when the client's problem is complicated. The quantitative parameters could assist them to understand the performance of operating the pointing device in various tasks. But the experts have no idea about the impact resulted from the client's motor control problem if there are no video. This integrated system could not only provide the quantitative parameters but the visualized process of operating the device to complete the task.

Acknowledgment. The authors would like to thank the National Science Council of the Republic of China for financially supporting this research under Contract No. NSC 95-2524-S-415-001-

References

1. Hwang, F., Keates, S., Langdon, P., Clarkson, J.: Mouse Movements of Motion-Impaired Users: A Submovement Analysis. In: Proceeding of ASSERS'04, pp. 102–109 (2004)
2. Keates, S., Hwang, F., Langdon, P., Clarkson, P.J.: The Use of Cursor Measures for Motion-Impaired Computer Users. Universal Access in the Information Society 2(1), 18–29 (2002)
3. Mackenzie, I.S., Kauppinen, T., Silfverberg, M.: Accuracy Measures for Evaluating Computer Pointing Devices. In: Proceeding of CHI, pp. 9–15 (2001)
4. Mithal, A.K., Douglas, S.A.: Difference in Mouse Microstructure of the Mouse and the Finger-Controlled Isometric Joystick. In: Proceeding of CHI, pp. 13–18 (1996)
5. Phillips, J.G., Triggs, T.J.: Characteristics of Cursor Trajectories controlled by the Computer Mouse. Ergonomics 44(5), 527–536 (2001)
6. Soukoreff, R.W., MacKenzie, I.S.: Towards a Standard for Pointing Device Evaluation, Perspectives on 27 Years of Fitts Law Research in HCI. Int. J. Human-Computer Studies 61, 751–789 (2004)
7. Chen, M.C., Chu, C.N., Wu, T.F., Yeh, C.C.: Computerized Assessment Approach for Evaluating Computer Interaction Performance. In: Miesenberger, K., Klaus, J., Zagler, W., Karshmer, A.I. (eds.) ICCHP 2006. LNCS, vol. 4061, pp. 450–456. Springer, Heidelberg (2006)
8. Zhou, H., Hu, H.: A Survey-Human Movement Tracking and Stroke Rehabilitation. University of Essex, Colchester United Kingdom (2004)
9. Hourcade, J.P., Bederson, B.B., Druin, A., Guimbretiere, F.: Differences in Pointing Task Performance Between Preschool Children and Adults Using Mice. ACM Transaction on Computer-Human Interaction 11(4), 357–386 (2004)

The BIC Model: A Blueprint for the Communicator

J.P. de Ruiter and N.J. Enfield

Max Planck Institute for Psycholinguistics

Abstract. In this paper, we outline a cognitive architecture for communicators, called the BIC model. The model consist of three main components. First, a (B)iological component, in which the genetic or built-in capacities of the communicator are specified. Second, an (I)nteraction Engine which uses neo-Gricean mutual simulation to attribute communicative intentions to signals, and to create signals to convey communicative intentions to the I-system of other agents. The third component of the BIC model is a content addressable database of (C)onventions which is used to store form/meaning mappings that have been successfully computed by the I-system. These stored form/meaning mappings are indexed by types of communicative context, so they can be retrieved by the I-system to save computational resources. The model can be used both as a computational architecture for a communication module in an artificial agent and as a conceptual model of the human communicator.

1 Introduction

Claude Shannon's [1] influential treatment "A mathematical model of communication" still influences the way many cognitive scientists think about communication. However, Shannon's theory is a theory of *data transmission*, not of real communication between two autonomous agents. The point of his work was to show how a series of bits could reliably be transmitted over a channel of limited capacity, and how this data transmission could be made less susceptible to the main enemy of communication engineers: noise (see Figure 1).

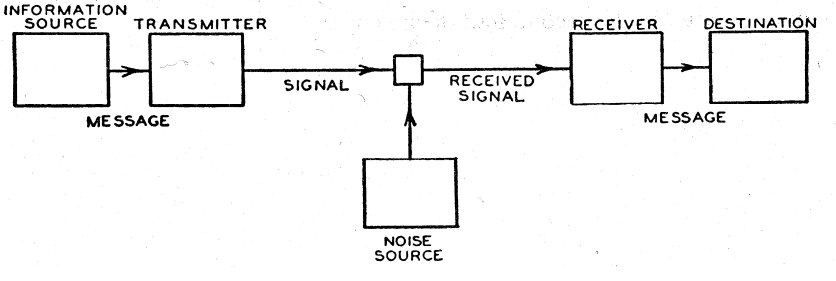


Fig. 1. Shannon's (1948) famous diagram

When applied to communication between agents (either human or artificial) the transformations that take place at the transmitter and receiver's end take on a

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 251–258, 2007.
© Springer-Verlag Berlin Heidelberg 2007

Saussurean character: for successful communication, there has to be an unambiguous mapping of signals (e.g., bit streams or speech sounds) to meanings, and a mapping of meanings onto signals that is its inverse. In psycholinguistic models of communication, this assumption is reflected at the level of phonology, syntax, and semantics (see e.g. [2]). In artificial communicators this assumption still guides implementations of speech recognition, comprehension, and generation. These levels of representation enable the communicator to unambiguously map signals onto meanings and vice versa. In this framework, the only thing that can go wrong is - as in Shannon's theory - noise that disrupts the transmission of symbols (e.g. phonemes) which then obviously also disrupts the subsequent interpretation of the signal, because if the phonemes are wrongly recognized, the words and sentences will be wrongly recognized as well. While this is indeed one of the challenges any communicator has to meet, it is not the only challenge, and probably not the biggest one either. Indeed, having a bad phone connection when we talk to other humans, or having too much ambient noise for automatic speech recognition to operate reliably is disruptive to communication.

A problem that is much harder to cope with is the *pragmatics problem*: in human communication, there is no 1:1 mapping between signals and their intended meanings. This is because the meaning of a communicative signal is not entirely fixed and context free, but depends on the *communicative intention* of the generator of that signal in a particular context (i.e., given a particular speaker, addressee, physical setting, and discourse history). The linguistic discipline of pragmatics is concerned with exploring this problem and its possible solutions, and the pragmatics literature is littered with examples showing that the mapping of signals to communicative intention is not 1 to 1, but rather Many to Many [3-6]. That is to say, every signal can convey many different communicative intentions, and every communicative intention can be conveyed by generating many different signals. For example, the phrase "thank you" can be intended as the expression of gratitude (when accepting a gift), the expression of annoyance (e.g. ironically, after someone has just dropped his coffee in your lap), or even as a request (a tennis umpire trying to quieten down the spectators during a match). However, the tennis umpire can also try to quieten down the spectators be saying "please be quiet", "quiet please", or "shhhhhht!". To complicate things further, communicators not only use language to communicate, but also facial expressions, gestures, eye-gaze, and functionally motivated actions (e.g. placing the book one wants to buy on the counter of the shop). For these nonverbal signals, it is even harder to establish an unambiguous meaning associated with it, as there are no "dictionaries" to guide us.

In this paper, we outline a blueprint (which we call 'the BIC model') for a real communication module, one that incorporates the pragmatics problem by making it central. The model is intended to serve as an architectural guideline for communication modules in artificial systems, but is also sufficiently general to use as a cognitive-psychological model of communication between any possible combination of humans, animals, and machines. The core assumption of the model is that for real communication, there are three distinct and essential components needed: a set of "biological" or "built in" constraints (the B-system), an interaction engine or social-interactional cognition component along the lines of [7] (the I-system), and a

content-addressable database containing conventions (the C-system). We now discuss the components of the BIC model, its implementation and some of its applications.

2 Biological or Built-in Constraints

Whether biological or artificial, an agent that communicates has physical constraints that limit the type of signals that it can produce and perceive. Robots have cameras with a certain resolution and sampling rate, humans have eyes with specific perceptual limitations (e.g. the visible colors). Robot use microphones, while humans have ears. Both can detect only a limited range of frequencies. These constraints limit the possible signals, with consequences for the way communication proceeds. In humans, the B-system also represents whatever communicative intelligence is innate, and is responsible for the human infant to start 'bootstrapping' into communication (e.g. by responding to certain frequency patterns, recognizing its mother's face, etc.). This is necessary, as the C(onvention)-system, discussed below, is assumed to still be empty in the newborn infant. In a communication device a la Shannon (e.g. a computer modem) the B-system is the capacity to distinguish 1s and 0s, leading to the capability to receive a sequence of 'bits'. The same holds for production: a robot has a number of output devices (a speaker, a head, a face, hands, etc) just as humans have faces, a voice, eyes etc. Even in the simple case of the Shannon device, a modem should be capable of creating 0s and 1s to send messages to the other side.

The anatomy of the human body, including its articulatory organs, is subject to constraints that are determined by the genetic make-up of homo sapiens. In the case of Robots, these constraints are in its physical design, and ideally these are as similar as possible to the human constraints, but this is generally not the case. While a range of artificial sensors exist that are more sensitive and accurate than the human receptors, there are as yet no artificial effectors more dextrous and flexible than the human hand.

3 The Interaction Engine

The Interaction Engine, or I-system, is a processor that is central in the BIC model. It has three main tasks. First, it performs the computations necessary to attribute a communicative intention to the sender of received signals (delivered to it via the B-system). Second, it performs the computations necessary to generate signals that will enable others' I-systems to successfully attribute communicative intentions. These two functions are essential, because of the pragmatics problem mentioned in the introduction: the Many to Many relationship of signals to communicative intentions. A third and important additional function of the I-system is to manage the C-system (the database of stored conventions), which will be discussed below. Importantly, the I-system is modality-independent; it will process a verbal utterance, the raising of an eyebrow, or gestures that accompany speech (see [8]) in the same way.

There have been different proposals about how to design an I-system. In the dialogue processing model of Pickering & Garrod [2] the I-system is assumed to be unnecessary most of the time. It only kicks into action when special conditions apply,

such as persistent misunderstandings or deliberate deception. According to Pickering and Garrod, the reflexive computations that are presumably involved in solving the pragmatics problem are computationally too expensive to be constantly governing ordinary communication.

In the approach to the pragmatics problem called *relevance theory* by Sperber & Wilson [6] it is assumed that the correct form/meaning mapping can be extracted by giving precedence to inferences that maximize 'relevance' while minimizing the cognitive effort needed to compute it. In contrast to Pickering and Garrod, Sperber and Wilson do claim that some type of inferential processing is necessary for communication, but they differ from the Gricean approach [3, 9] in the way this processing is performed. They assume that Grice's maxim of Relevance is central, and sufficient.

The approach we take for building an I-system is essentially neo-Gricean, following Levinson [5, 7]. We propose that the I-system solves the pragmatics problem by using conceptual *simulation* of the interlocutor(s). In short, senders of signals simulate how a receiver would interpret the signal, and receivers simulate how a sender could have decided to send that particular signal. Recent neurocognitive findings from the EU funded project JAST have shown that communicators solving the pragmatics problem while sending signals activate the same well-identified brain region as receivers do while solving the pragmatics problem for comprehension [10]. We call this type of simulation *conceptual* to distinguish it from *sensori-motor* simulation, as has been proposed by researchers who study "mirror neurons" [11][1]. In a conceptual simulation, what is simulated are the *inferences* of the interlocutor, given a certain set of propositions about the world (i.e., his beliefs and desires, see e.g. [12]) and a certain signal that is received or being sent. Such a simulation can never be perfect, as information about the beliefs and desires of others is always incomplete. However, propositions of which the truth value is both unknown and essential for the inferences at hand are assumed to be copied from the communicator's own beliefs and desires. (In any case, it is not critical to have an exact copy of your interlocutor's mental state, but to have one that is 'sufficient for current purposes'; Clark [13]) This is a computational *heuristic* that leads to a degree of 'egocentrism' in communication. Evidence suggests that communicators often do not take current common ground among interlocutors into account, for instance in producing and understanding referring expressions [14-17]. While communication proceeds, participants gradually build up a more and more accurate representation of the other (or, to put it in other words, they will develop more and more *common ground*, see [18]) which enables them to take their interlocutor's perspective into account. This process is beautifully illustrated in an experiment by Clark & Wilkes-Gibbs [19], in which communicators are shown to gradually establish so called *conceptual pacts*, which are essentially form/meaning mappings that have been implicitly agreed upon by communicators. Updating our model of interlocutors during the course of communication has been called *grounding* by Clark & Schaefer [20].

While we propose to build an I-system based on mutual simulation, we appreciate the point by [14-17] and others that reflexive computations can be expensive. To

[1] In the fMRI study reported in De Ruiter et al., there was no activation in the mirror neuron areas while communicators were trying to solve the pragmatics problem.

avoid this resource bottleneck, the I-system can store the results of its pragmatic simulations in the C-system, which constitutes a content-addressable database of recurrent form/meaning mappings. Take for example the following short dialogue:

A: What did you think of Steve's latest book?
B: It put me to sleep.

The first time A encounters B's response, she might well need to engage in conceptual simulation to extract the communicative intention from the speech "It put me to sleep" (as the 'literal' meaning does not make much sense), but after having inferred (and validated, through grounding) the communicative intention of the utterance, the mapping of the signal "X put me to sleep" onto the communicative intention "I want you to know that I found X rather boring" can be stored for later use. Thus, we now discuss the third and final component of the BIC model, the C-system.

4 Conventions

Efficient communication between agents is impossible without the ability to rely on socially shared *conventions* about what signals mean. For example, most if not all implemented models for linguistic communication contain a lexicon, a database that stores the relationship between the surface form of words and their associated meanings. However, storing conventional form/meaning mappings need not be limited to words. It can also be used to store mappings between phrases and their associated communicative intention. For example, the phrase "How are you?" is frequently not a request for information, but a greeting, and storing this mapping can save the I-system a lot of processing cycles. Similarly, parameterized *templates* could be stored as well. The form "Could you <VP>?" is most often a polite request to VP. Again, storing this mapping avoids having the I-system compute the communicative intention from scratch every time it is encountered.

Syntactic rules also establish a system of shared rules that govern the ordering of and mutual relations between linguistic symbols. At a more peripheral level of processing, the same holds for phonology. In fact, traditional linguistics has mainly focused on charting out the conventions governing linguistic competence. However, we see the C-system as a place to store not only the general linguistic *competence* of the language user, but also aspects related to performance, or language use. For instance, the *conceptual pacts* studied by Clark & Wilkes-Gibbs [19] can be treated as local conventions, conventions that only apply for a certain context and with a certain interlocutor. A good example of local conventions can be found in the domain of person reference [21]. In a limited social group where everybody is assumed to know him, an individual can be referred to as "Herb", but in the wider context of a psycholinguistics conference, the reference "Professor Clark" might be needed to achieve recognition. The localized nature of conventions represents a challenge for the implementation of the C-system. Conventions have a "social scope" which needs not only to be represented in the C-system, but in fact provides an essential *retrieval key*. If I want to refer to a person, I do not want to look through all the social groups I am a part of in sequence before deciding on a reference: I want to be able to index the retrieval of the reference by the social group or individual I am designing my person

reference for. A similar argument can be made for the *context* of communication. When I'm socializing in a bar with my friends, other conventions apply than when I am in a formal meeting at work, even if I am in that meeting with the same people. By using both the social scope and the social context as a retrieval cue for the C-system, the I-system can retrieve those conventions that are 'relevant' to the situation at hand first, before considering other conventions.

To summarize, when the I-system receives a signal via the receptors of the B-system, it first checks the C-system to see if that signal with that context (or parts thereof) has been stored. If it has, the I-system will use that stored mapping as a basis for attributing a communicative intention to the signal. If it is not, the I-system will engage in simulation of the interlocutor, and try to compute the communicative intention. If this turned out to be the correct interpretation, it will store the newly found mapping into the C-system. Analogously, if the I-system needs to produce a signal to get across a certain communicative intention, it will first look in the C-system to see if it contains a previously stored meaning-form mapping that can be used. If not, it will engage in simulation (of the receiver) to generate a signal that might work. If it turns out to have worked, this new mapping will also be stored in the C-system for later work. To optimize retrieval efficiency, the C-system is a content-addressable database that allows social scope and context to be used as retrieval keys.

5 Summary

In this paper we have outlined a blueprint for a communication module (see Figure 2). It is based on three essential components: (1) a set of biological (hard-wired) constraints that define the properties of the possible incoming and outgoing signals, called the B-system, (2) an interaction engine that performs the inferential processing

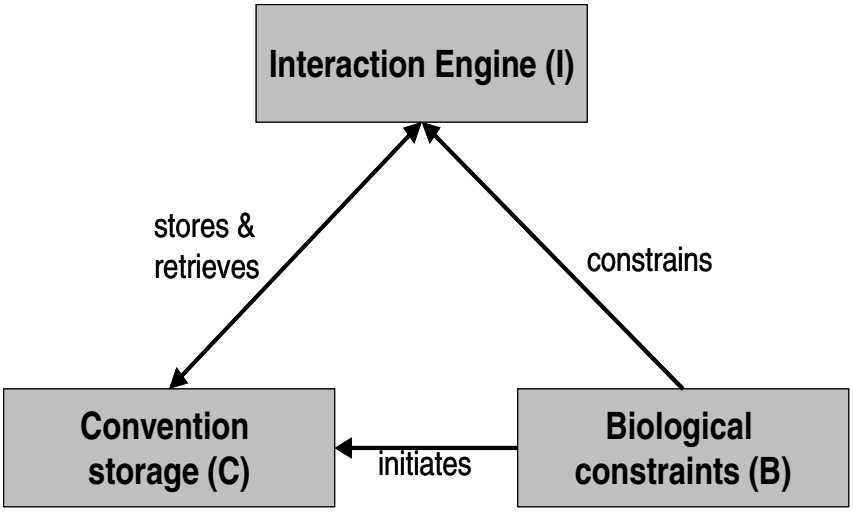


Fig. 2. General layout of the BIC architecture for agent-agent communication

needed to solve the pragmatics problem of mapping communicative intentions to signals and vice versa, and (3) a database of socially shared communicative conventions about the mappings between signals and communicative intentions. The more "experience" the system will have, the more sophisticated and efficient it will operate. For the implementation of the I-system, we propose to use neo-Gricean mutual simulation, combined with rational use of the C-system to take shortcuts whenever possible.

The BIC model can be used both as a conceptual model in the study of agent-agent communication (including HCI), and as a blueprint for an implementation of a communication module in artificial agents. This way, insights about communication between agents can flow in both directions. Challenges arising from implementing a communication module in agents can stimulate interesting new research issues in human-human communication, and existing knowledge of the latter can be used to improve the quality of our implementations in artificial agents.

References

1. Shannon, C.E.: A mathematical theory of communication. Bell. System Technical Journal 27, 379–423 (1948)
2. Pickering, M.J., Garrod, S.: Toward a mechanistic psychology of dialogue. Behavioral and Brain Sciences 27(2), 169–226 (2004)
3. Grice, H.P.: Logic and conversation, In: Cole, P., Morgan, J.L. (eds.) Syntax and semantic 3: Speech acts, pp. 41–58. Academic press, New York (1975)
4. Levinson, S.C.: Pragmatics. Cambridge University Press, Cambridge (1983)
5. Levinson, S.C.: nteraction biases in human thinking. In: Goody, E.N. (ed.) Social intelligence and interaction, pp. 221–260. Cambridge University Press, Cambridge (1995)
6. Sperber, D., Wilson, D.: Relevance: communication and cognition, 2nd edn. Blackwell, Oxford, UK (1995)
7. Levinson, S.C., On the human interactional engine. In: Enfield, N.J., Levinson, S.C. (eds.) Roots of Human Sociality: Culture Cognition, and Interaction,Berg: London (2006)
8. De Ruiter, J.P.: Postcards from the mind: the relationship between thought, imagistic gesture, and speech. Gesture 7(1), 21–38 (2007)
9. Levinson, S.C.: Presumptive Meanings; The Theory of Generalized Conversational Implicature. The MIT Press, Cambridge, Massachusetts (2000)
10. De Ruiter, J.P., et al.: On the origin of intentions, in Attention & Performance XXII. (in press)
11. Gallese, V., Goldman, A.: Mirror neurons and the simulation theory of mind reading. Trends in Cognitive Science 12, 493–501 (1998)
12. Jurafsky, D.: Pragmatics and comutational Linguistics. In: Horn, L.R., Ward, G. (eds.) The Handbook of Pragmatics, pp. 578–604. Blackwell, Oxford (2006)
13. Clark, H.H.: Using language. Cambridge University Press, Cambridge (1996)
14. Bard, E.G., et al.: Sharing the cognitive burdens of dialogue. (in press)
15. Barr, D.J., Keysar, B.: The paradox of egocentrism in language use. In: Colston, H.L., Katz, A.N. (eds.) Figurative language comprehension: Social and cultural influences, pp. 21–42. Erlbaum, Mahwah, N.J (2005)
16. Keysar, B., Barr, D.J., Balin, J.A.: Definite Reference and Mutual Knowledge: Process Models of Common Ground in Comprehension. Journal of Memory and Language 39, 1–20 (1998)

17. Horton, W.S., Keysar, B.: When do speakers take into account common ground? Cognition 59, 91–117 (1996)
18. Clark, H.H., Marshall, C.R.: Definite reference and mutual knowledge. In: Joshe, A.K., Webber, B.L., Sag, I.A. (eds.) Elements of discourse understanding, pp. 10–63. Cambridge University Press, Cambridge, U.K (1981)
19. Clark, H.H., Wilkes-Gibbs, D.: Referring as a collaborative process. Cognition 22, 1–39 (1986)
20. Clark, H.H., Schaefer, E.F.: Contributing to discourse. Cognitive Science 13, 259–294 (1989)
21. Enfield, N.J., Stivers, T. (eds.): Person reference in interaction: linguistic, cultural and social perspectives. Cambridge University Press, Cambridge (2007)

Conveying Browsing Context Through Audio on Digital Talking Books

Carlos Duarte and Luís Carriço

LaSIGE – Faculty of Sciences of the University of Lisbon
Edifício C6, Piso 3, Campo Grande
1749-016 Lisboa, Portugal
{cad,lmc}@di.fc.ul.pt

Abstract. This paper presents the results of a study comparing the use of auditory icons, earcons and speech in an audio only interface for a digital talking book player. The different techniques were evaluated according to the identification errors made, and subjective measures of understandability, intrusiveness and pleasurability. Results suggest the use of auditory icons combined with speech whenever necessary, in detriment to the use of earcons, for applications sharing the characteristics of digital talking book players.

Keywords: Evaluation, Audio Interfaces, Auditory Icons, Earcons, Speech, Digital Talking Books.

1 Introduction

Digital recordings of book narrations synchronized with their textual counterpart allow for the development of digital talking books (DTB), supporting advanced navigation and searching capabilities, with the potential to improve the book reading experience for visually impaired users. By introducing the possibility to present, using different output media, the different elements comprising a book (text, tables, images) we reach the notion of Rich Digital Book [1]. These books, in addition to presenting visually or audibly the book's textual content, also present the other elements, and offer support for creating and reading annotations.

Current DTB players do not explore all the possibilities that the DTB format offer. The more advanced players are executed on PC platforms, and require visual interaction for all but the most basic operations, behaving like screen readers, and defeating the purpose to serve blind users [2]. We have developed an adaptive multimodal player, supporting visual and audio interaction, able to adapt its interaction to user characteristics [3]. In this paper we detail the approach taken in the development of the player features that specifically target blind users.

The DTB format, possessing similarities with HTML, has, nevertheless, some advantages from an application building perspective. The most important one is the complete separation of document structure from presentation. Presentation is completely handled by the player, and absent from the digital book document. Navigation wise, the user should be able to move freely inside the book, and access its

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 259–268, 2007.
© Springer-Verlag Berlin Heidelberg 2007

content at a fine level of detail. The table of contents should also be navigable. One major difference between a DTB player and a HTML browser is the support offered for annotating content. Mechanisms to prevent the reader becoming lost inside the book, and to raise awareness to the presence of annotations and other elements, like images, are also needed.

To solve these problems in an audio only environment (speech recognition plus auditory display) several approaches have been tested. Concerning the auditory display, playback of pre-recorded books is complemented with three other solutions: pre-recorded speech cues, auditory icons [4] and earcons [5]. These solutions are used to convey context information and navigational cues.

This paper presents the evaluation of several approaches to the problem of how to convey the current context while browsing a rich DTB. The following section details the different book presentation elements, and the approaches that have been tested for conveying their presence. Section 3 presents the experimental setting used in the evaluation of the different approaches. The evaluation results are given in section 4, and section 5 discusses these results, and derives a set of guidelines for developing applications with characteristics similar to those of a DTB player.

2 Audio Presentation Techniques in Rich Digital Talking Books

DTBs are capable of presenting their contents either on screen or through speech, recorded or synthesized. Besides the main content presentation, other book elements also have to be presented when working in an audio only environment. This means that the table of contents and the annotation must have an audio representation also. If the annotation is a voice annotation this is straightforward. If it is a text annotation, its content can be reproduced using a speech synthesizer.

However, in an audio only environment, not only content has to be transmitted, but the entire narration context has to be available in an audible format, thus enabling the reader to form an accurate image of the surrounding elements. For this to be possible the reader must be aware of annotations and images present in the book, as well as be able to know what is her/his position in the book whenever desired.

To understand how to better transmit this information to the reader, we evaluated three techniques for improving the reader awareness to the different DTB elements: speech, auditory icons and earcons.

Using speech for transmitting context information is perhaps the easiest of the three approaches, involving just the selection and recording or synthesis of the words to employ. While for certain applications this may not be a trivial task, in the DTB context, where the elements are well identified, it is an uncomplicated one. Speech can also be expected to be the technique where the message meaning is most easily understandable by the listener.

However, the use of speech can have disadvantages also. Since the book's content is being narrated, there will be two audio tracks presenting information in the same way. If the presented messages are long they can disrupt the reading experience, become too intrusive, or even make it harder to listen to the main content if both tracks are played back simultaneously [6]. Furthermore, for voice messages to be understood, the listener must know the language in which the messages are spoken.

Auditory icons have been defined by Gaver [7] as "*Everyday sounds mapped to computer events by analogy with everyday sound-producing events*". Due to this nature, auditory icons share with voice commands the ease of understanding, if enough care is put into the auditory icons selection, ensuring appropriate and intuitive mappings between the sounds and what they represent in the interface. However, there may be cases, where it may be difficult, and even impossible, to find a sound to map to abstract interface events or components [8]. In the DTB domain, certain concepts are abstract enough to make it harder to find an everyday sound to map to, e.g. the beginning of a chapter.

Earcons are "*abstract, synthetic tones that can be used in structured combinations to create auditory messages*" [9]. They can be used in the situations where there are no intuitive sound to represent an interface's event. This gives them the advantage of being able to represent any event or interaction with the interface. They are based on an abstract mapping between a music-like sound and the interface events, which means that, at least initially, they have to be explicitly learned.

There are four types of earcons [5]: one-element, compound, hierarchical and transformational, allowing them to be used in every situation, and even giving them the flexibility to be concatenated, in a process similar to building sentences out of words [9]. Guidelines on how to build earcons are also available [10], identifying timbre, rhythm, pitch and register as sound characteristics that can be used to effectively differentiate one earcon from the others.

3 Experimental Setting

In order to understand what solutions are more appropriate for the different DTB elements, and how they can be used, an experiment was set up, evaluating the use of the three different techniques, in a purely audio version of the DTB interface. To better focus on this goal, we decided not to use the current version of the DTB player, instead conducting a Wizard of Oz evaluation, with just the features required for an audio environment.

Four elements, essential for contextual awareness, were the subject of evaluation: beginning of a new chapter, current chapter number, presence of an annotation and presence of an image. A pre-recorded narration of the book "O Senhor Ventura" by a professional narrator was used in the experiment. Four excerpts of the narration, each making use of different audio feedback techniques, were prepared:

1. The first excerpt, six minutes and 39 seconds long, consisted of four chapters. Chapter beginnings, the presence of annotations, and the presence of images were signaled with earcons. The current chapter number was transmitted by speech recordings. When listening to the chapter number, the book's narration was paused.
2. The second excerpt, seven minutes and 38 seconds long, consisted of three chapters. Chapter beginnings were announced by a speech recording of the chapter number. Speech was also used to signal the presence of images and annotations. User requests for chapter numbers were answered with earcons, with no interruption of the book's narration.

3. The third excerpt, six minutes and 36 seconds long, consisted of three chapters. Chapter beginnings were announced with an earcon. Auditory icons signaled the presence of images and annotations. Speech recordings were used to transmit the chapter number, without pausing the book's narration.
4. The fourth and last excerpt, six minutes and three seconds long, consisted of three chapters. All feedback was given with earcons. The chapter number announcements paused the book's narration.

The speech used consisted of pre-recordings of the words "*annotation*", "*image*", and of the chapter numbers. Each chapter number was recorded on its own, meaning that there were no composition of recordings of individual numerals.

Auditory icons were used to signal the presence of images and annotations. For the image signal, the sound a photographic camera shutter closing was employed. The sound's duration was 600 milliseconds. For annotations, the sound of a typewriter was used. This sound's duration was 3 seconds and 50 milliseconds. This last sound was larger than the first because it was expected to be more difficult to recognize.

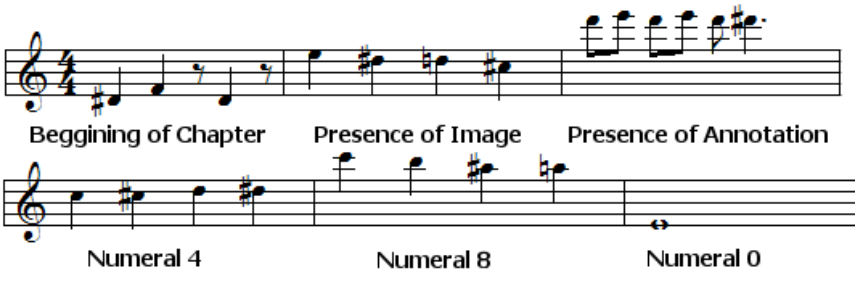


Fig. 1. Earcons: Beginning of chapter, presence of image and annotation, and three examples of numeral used for chapter numbers

Earcons were designed to signal chapter beginnings, presence of images and annotations, and chapter numbers. Figure 1 presents the earcons used in the evaluation procedure. To promote ease of identification, each earcon is designed according to the earcon design guidelines [10]. Different timbres are employed: chapter beginnings – marimba; images – synth bass; annotations – trombone; numerals 1 to 4 – acoustic piano; numerals 5 to 9 – organ; and numeral 0 – tubular bells. The earcons for chapter numbers were divided into three groups corresponding to numerals 1 to 4, numerals 5 to 9 and the numeral zero. The numerals 1 to 4 are played with the same timbre, each numeral consisting of one more note than the previous, played in an ascendant scale. The numerals 5 to 9 are played with a different timbre, following the same principles, but with each note played in a descendant scale. The numeral zero is played with yet another timbre. Numbers above 9 were composed with sequential presentation of the individual numerals (e.g. the number 15 is presented by playing the numeral 1 followed by the numeral 5). The interval used between numerals when composing numbers was 400 milliseconds.

3.1 Procedure

Seven participants aged between 21 and 26, one female and six males, undertook the experiment. The experiment was a within-participants factorial design with two independent variables. The first independent variable was the type of auditory feedback technique used. The second variable defined how the current chapter number was presented: with or without interruption of the main narration. Dependent variables were the number of correct identifications of book elements, and subjective measures of understandability, intrusiveness and pleasurability. The main hypothesis was that varying the type of auditory feedback used would lead to different levels of understandability, intrusiveness and satisfaction.

The experiment began with the presentation phase, where the participants were introduced to the different auditory feedback techniques. In this phase the participants were asked to recognize the sounds used as auditory icons, and to associate them with one of the features of the DTB player. This was followed by the presentation of the different earcons, repeated as many times as wished. When the participants felt comfortable with the earcons, these were played back twice in a different order, to test the recall rate. The construction of numbers from the numeral earcons was then explained to the participants. The participants were then asked to identify twelve numbers represented by earcons. This phase ended with the replay of earcons used for beginning of chapter, images and annotations.

The testing phase consisted in the presentation of the four book excerpts. During the excerpts presentation, participants were allowed to use three commands: pause and play, for controlling playback (no forward or backward movement was allowed) and another command to inquire the current reading position. The participants were asked to perform two tasks during excerpt presentation: one task consisted in keeping count of the number of annotations (or images – this varied from excerpt to excerpt) in that excerpt; the second task consisted in identifying, for all occurrences of images (or annotations), the current chapter number, writing it down and delivering it to the test coordinator, immediately after recognizing the audio cue. After each excerpt, the participants answered a questionnaire, rating the techniques used in the excerpt in terms of their understandability, intrusiveness, and pleasurability. Rating scales ranged from zero to nine, with zero meaning it was hard to understand the sound's meaning, the sound was very intrusive, and unpleasant. Nine corresponded to a sound with an easily identifiable meaning, not intrusive, and pleasant to listen to.

4 Results

The preparation phase allowed for the individual evaluation of the auditory icons and earcons, while their use as part of an application was evaluated during the next phase.

The auditory icons were correctly identified by all the participants. The sound of the camera shutter closing was associated with the image element by all participants. The typewriter sound was associated with the annotation element by six participants. The other participant associated the sound with the image element.

The three earcons for chapter beginning, images and annotations, presented twice to each participant, were correctly identified by just three participants. One participant

was not able to correctly interpret the chapter beginning and annotations earcons in the first round, exchanging their meanings. Two participants exchanged the meanings of the annotations and images earcons in both rounds of presentation. The other participant exchanged the meanings of the beginning of chapter and images earcons in both rounds. No clear misinterpretation pattern was identified. It is possible that all these misinterpretations are due to the participants not having heard the earcons enough times to correctly recall them.

The twelve number earcons for the number identification task represented the numbers 62, 8, 17, 2, 46, 93, 2, 30, 11, 54, 66, 9. Four were single digit numbers, six were composed by earcons of different timbres, and two by earcons of the same timbre. Three participants correctly identified all numbers (the same participants that had correctly identified all the earcons previously). Two participants incorrectly identified two numbers, and the other two participants incorrectly identified five numbers. The fourteen errors, out of the 84 numbers played, can be divided in the following categories: wrong count of notes in a numeral (e.g. identifying a three when a two as played) – 8; wrong association of timbre to numeral (e.g. identifying a two when a six was played) – 3; wrong interpretation of the pause between two notes (e.g. identifying a two when an eleven was played) – 3.The total percentage of correctly identified numbers was 83.3%.

4.1 Testing Phase Results

The four excerpts of the book played back to the seven participants contained a total of 385 fixed audio cues divided in the following way: 210 in the form of earcons, 84 in the form of auditory icons and 91 in the form of spoken messages. We will use this number of fixed audio cues as the corpus for comparison between the different techniques. To arrive at the total number of audio cues, the number of times the chapter number was requested would have to be added.

When considering the identification of audio cues, we expected that both spech and auditory icons would be identified correctly every time. This was indeed the case, with all participants identifying correctly all the elements when presented with these two techniques. When the elements were presented by earcons, the recognition rate lowered to 89.05%., corresponding to a total of 23 misinterpreted earcons over the 7 experiments. The percentages of incorrect interpretations by book element were as follows: 15.71% for the beginning of chapter earcon, 14.29% for the presence of images earcon, and 2.86% for the presence of annotations earcon. This might lead to believe that the beginning of chapter earcon and the presence of images earcons can be misinterpreted one for the other. We can further detail the analysis by looking at how the participants were interpreting the earcons. All the incorrectly identified presence of images earcons were mistaken for presence of annotations, and 72.73% of the beginning of chapter earcons were mistaken for presence of annotations (18.18% were mistaken for numerals and 9.09% for presence of images earcons). These results reveal that the beginning of chapter and the presence of images earcons are not being mistaken one for the other, but are being interpreted as presence of annotations earcons. This is somewhat surprising, since the presence of annotations earcon was correctly identified 97.14% of the times it was played.

The next results report the subjective measures obtained from the questionnaires. The first measure, understandability, can be expected to have a similar outcome to the identification results presented above. Thus, we expected higher values of understandability for speech and auditory icons than for earcons. Figure 2 presents the average understandability for the four elements, by type of audio cue used.

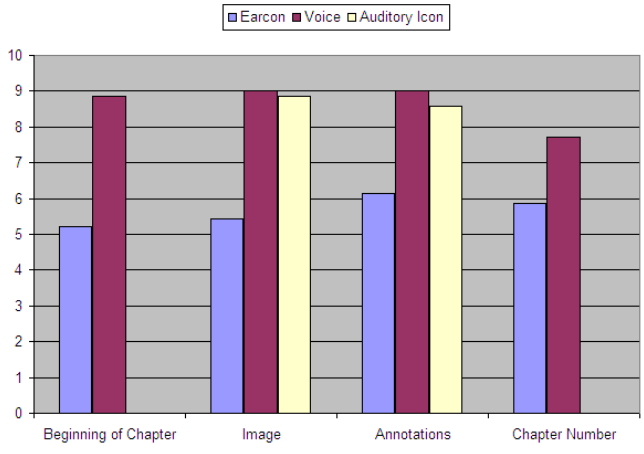


Fig. 2. Understandability of the different auditory cues used

To determine if the differences shown are statistically significant two t-tests (one for the beginning of chapter and chapter numbers auditory cues) and two ANOVA tests (one for the presence of images and other for the presence of annotations) were carried out. The t-test comparing the beginning of chapter results found a significant increase ($t(19) = 3.55$, $p < 0.01$) in understandability when speech was used instead of earcons. The t-test comparing the results for chapter numbers between earcons and speech also revealed a significant increase in understandability ($t(26) = 3.03$, $p < 0.01$). The ANOVA for the presence of image cues between earcons, speech and auditory icons was also found to be significant ($F(2, 18) = 40.98$, $p < 0.001$). Post hoc Tukey HSD tests found earcons to be have significant lower understandability than speech (HSD = 11.31, $p < 0.01$) and than auditory icons (HSD = 10.85, $p < 0.01$), and no difference between speech and auditory icons. The ANOVA for the presence of annotations understandability when using earcons, speech and auditory icons was also significant ($F(2, 18) = 10.47$, $p < 0.001$). Once again, post hoc Tukey HSD test showed that earcons had significant lower understandability than speech (HSD = 6.00, $p < 0.01$) and auditory icons (HSD = 5.10, $p < 0.01$). No significant difference was found between speech and auditory icons.

Figure 3 presents the average results for the intrusion rating of the three auditory cues employed (higher values mean less intrusive sounds). Once again, two t-tests and two ANOVA tests were performed to determine if the differences are statistically significant. The t-tests for the beginning of chapter and chapter number comparisons did not identify any significant results. The ANOVA for the intrusiveness when presenting images comparing earcons, speech and auditory icons found a significant

difference ($F(2, 18) = 4.01$, $p < 0.05$). Post hoc Tukey HSD tests however did not find significant results between any pair of results. t-tests with the Bonferroni adjustment found that earcons were significantly more intrusive than auditory icons for signaling the presence of an image ($t(12) = 3.62$, $p < 0.05$). The ANOVA test for the presentation of annotations with earcons, speech and auditory icons found a statistically significant difference. The post hoc Tukey HSD tests identified once again that earcons were more intrusive than auditory icons (HSD = 4.56, $p < 0.05$).

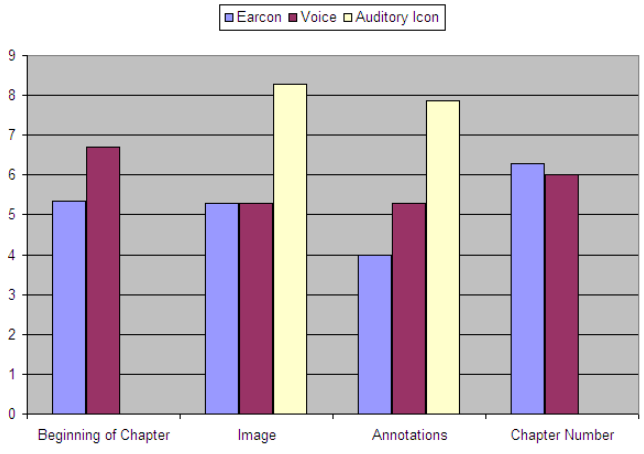


Fig. 3. Intrusion of the different auditory cues used. Higher values correspond to less intrusive sounds.

Figure 4 presents the average results for the pleasurability rating. The same t-tests and ANOVA tests were applied. The t-test for the beginning of chapter feedback revealed that participants found speech more pleasurable than the earcons ($t(19) = 3.28$, $p < 0.01$). Chapter numbers presented with speech were also found to be significantly more pleasurable than with earcons ($t(26) = 2.71$, $p < 0.05$). The ANOVA test for the presentation of image presence with earcons, speech and auditory icons found a significant difference ($F(2, 18) = 36.06$, $p < 0.001$). Post hoc Tukey HSD confirms that participants found earcons to be significantly less pleasurable than speech (HSD = 8.65, $p < 0.01$) and auditory icons (HSD = 11.54, $p < 0.01$). The corresponding ANOVA test for annotation presence signaling with earcons, speech and auditory icons also found a significant difference ($F(2, 18) = 13.67$, $p < 0.001$). Post hoc Tukey HSD tests once again confirmed that earcons were found to be significantly less pleasurable than speech (HSD = 5.43, $p < 0.01$) and auditory icons (HSD = 7.06, $p < 0.01$).

The effect of interrupting the narration of the main content when presenting the current chapter number on the subjective ratings was also studied. However no significant results were found for understandability, intrusion and pleasurability ratings, which points to the important factor for these ratings being the type of audio feedback technique employed.

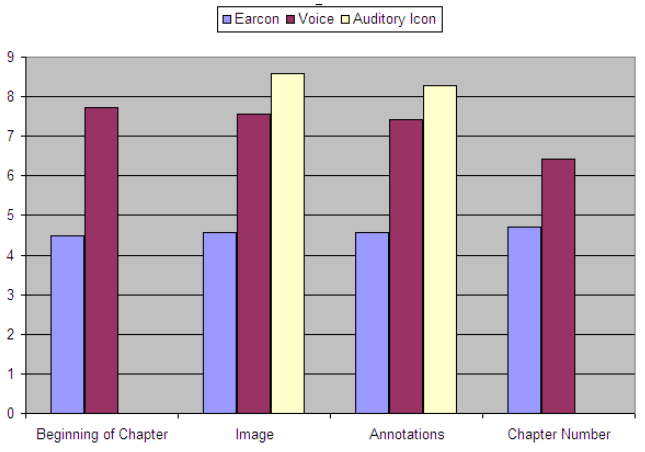


Fig. 4. Pleasurability of the different auditory cues used

5 Conclusions

The results presented in the previous section indicate that auditory icons and spoken messages should be preferred to earcons in the design of audio DTB players' interfaces. Earcons proved to be more prone to identification errors, and accordingly, test participants found them less suited to transmit the correct meaning. In addition, the results also show that participants found earcons the least pleasurable of all the evaluated techniques. When considering the intrusion results, earcons and speech achieve comparable results, but both techniques were considered significantly more intrusive than auditory icons.

Observations made during the experiments support these results. It was common amongst test participants to need more time to identify the meaning of a sound when presented with earcons. This is supported by the number of times most participants requested a pause in excerpts which made use of earcons to signal the presence of images or annotations, compared to other excerpts. Another evidence was the request for chapter numbers when they were presented using earcons in comparison with other techniques. Although some participants did the request just for confirmation (the correct number was already written down) it nevertheless shows that participants felt less secure with the earcons.

When comparing test performance on the first and last excerpts, which were the ones which relied most in earcons, all measures evolved positively with the exception of the understandability of the earcon for signaling the presence of an image (average of the answers dropped slightly from 5.43 to 5.29) and the intrusiveness for chapter beginnings, presence of images and annotations. The greatest evolutions were felt in the understandability and pleasurability of the presence of annotations and chapter numbers earcons. This may imply that with more time to familiarize with the earcons used, the measures could continue to evolve positively. However, one cannot be sure until further tests confirm this hypothesis.

For applications sharing the characteristics of a DTB player, we recommend the use of auditory icons and speech. As the events needing audio feedback might not occur frequently in this kind of applications, earcons are at a disadvantage, since it will be harder to memorize and associate their sound with an event, due to the mentioned low frequency of events. The events comprehension should also require the least amount of cognitive effort by the listener, since listening to the book content is the primary task. This is another factor that impacts negatively the use of earcons. We also suggest that auditory icons should be used whenever possible, due to normally being of shorter duration than speech messages. This means smaller interruptions of the book content narration. Another advantage of auditory icons is the fact that they are more universal than any language that may be used, thus requiring less effort for interface development. For the situations where it is difficult to find an auditory icon, then speech can be used to good effect.

References

1. Carriço, L., Duarte, C., Lopes, R., Rodrigues, M., Guimarães, N.: Building Rich User Interfaces for Digital Talking Books. In: Jacob, R.J.K., Limbourg, Q., Vanderdonckt, J. (eds.) Computer-Aided Design of User Interfaces IV, pp. 335–348. Springer, Heidelberg (2005)
2. Duarte, C., Carriço, L.: Users and Usage Driven Adaptation of Digital Talking Books. In: Proceedings of the 11th International Conference on Human-Computer Interaction, Las Vegas, Nevada, USA (2005)
3. Duarte, C., Carriço, L.: A conceptual framework for developing adaptive multimodal applications. In: Proceedings of the 11th International Conference on Intelligent User Interfaces, Sydney, Australia, pp. 132–139. ACM Press, New York (2006)
4. Gaver, W.W.: Auditory Icons: Using Sound in Computer Interfaces. Human-Computer Interaction, vol. 2(2), pp. 167–177. Lawrence Erlbaum Associates, Inc. Mahwah, NJ (1986)
5. Blattner, M., Sumikawa, D., Greenberg, R.: Earcons and Icons: Their Structure and Common Design Principles. Human-Computer Interaction, vol. 4(1), pp. 11–44. Lawrence Erlbaum Associates, Inc. Mahwah, NJ (1989)
6. Petrie, H., Johnson, V., Furner, S., Strothotte, T.: Design Lifecycles and wearable computers for users with disabilities. In: Proceedings of the First International Workshop of Human Computer Interaction with Mobile Devices, Glasgow, Scotland. Department of Computing Science, University of Glasgow (1998)
7. Gaver, W.W.: Auditory Interfaces. In: Helander, M.G., Landauer, T.K., Prabhu, P.V. (eds.) Handbook of Human-Computer Interaction, 2nd edn. vol. 1, pp. 1003–1041. Elsevier, Amsterdam (1997)
8. Brewster, S.A.: Overcoming the lack of screen space on mobile computers. Personal and Ubiquitous Computing, vol. 6(2), pp. 188–205. Springer, Heidelberg, New York (2002)
9. Brewster, S.A.: Providing a Structured Method for Integrating Non-Speech Audio into Human-Computer Interfaces. PhD Thesis, Department of Computer Science, University of York (1994)
10. Brewster, S.A., Wright, P.C., Edwards, A.D.N.: Experimentally derived guidelines for the creation of earcons. In: Proceedings of BCS-HCI, Huddersfield, UK, pp. 155–159. Springer, Heidelberg, New York (1995)

An RFID-Based Point-and-Listen Interface Providing Library Access for the Visually Impaired

Daniel Fallman, Kent Lindbergh, Oskar Fjellstrom,
Lars Johansson, Fredrik Nilbrink, and Linda Bogren

Umeå Design Research Group, Umeå Institute of Design,
Umeå University, SE-90187, Umeå, Sweden
{ daniel.fallman,kent.lindberg,oskar.fjellstrom,
lars.johansson,fredrik.nilbrink,linda.bogren}@dh.umu.se

Abstract. We present the AudioIndex prototype, a light-weight mobile RFID-based point-and-listen interface that allows visually impaired to browse and search for audio books within a public library without the need for library staff guidance. AudioIndex has been specifically designed to fit well into the library environment and its routines, a fit made possible by carefully considering interaction styles and by solving a number of technical design challenges, including distribution of data, hardware platform, and RFID specifications.

Keywords: Pointing Interface, RFID, Mobile, Wireless, Visually Impaired, Speech Synthesis, Design.

1 Introduction

The use of RFID (Radio Frequency Identification Tag) tags and readers has increased rapidly over the last years, showing ingenuity in terms of technology as well as in applications. Most RFID systems deployed in library environments however tend to focus on solving logistical issues, e.g. automated book check-outs and returns, designed to save time and money. In contrast to these efforts, the AudioIndex project presented in this paper is an attempt at using RFID for something else—as a means for improving library accessibility for visually impaired people. The AudioIndex prototype is a light-weight mobile RFID-based point-and-listen interface that allows visually impaired to browse and search for audio books within a public library without the need for library staff guidance.

While public access to library services on a high level can be regarded as essential for society, as literacy and free access to information are fundamental conditions for anyone seeking to participate in the democratic process, the incentives for this project comes from the fact that a number of people do not have access to libraries in the same way as most people do. A particularly vulnerable group is those that due to limitations in their cognitive or physical abilities have no or limited access to library services. Throughout this project, we have been working in close collaboration with end users and representatives from the Swedish Disability Federation, and their

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 269–278, 2007.
© Springer-Verlag Berlin Heidelberg 2007

feedback has continuously been used in the implementation process to guide design decisions.

2 Project Overview

The Audio Index prototype system has been implemented in eight copies that are currently in operation and undergoing long-term public use and testing at Umeå City Library, Umeå, Sweden. In this paper, we will mainly focus on various aspects of the technical implementation of these prototypes. The prototypes are the result of a larger accessibility design project that has been carried out in four stages between 2002 and 2006, described briefly below. For a more detailed description of this process, see [4].

In the first phase, we compiled existing guidelines for improving accessibility for the visually impaired and carried out field studies and interviews. Based on the information thus gathered, we provided recommendations to the library for enhanced accessibility to its services. These recommendations were specifically tailored to cater to the needs of people with various kinds of cognitive and physical disabilities but would in fact improve library accessibility for all library visitors.

In a second phase, we employed creative group methods and techniques such as brainstorming to generate ideas for increased accessibility to libraries especially for visually impaired visitors. These efforts resulted in a final concept and use scenario—which the prototype implementation reported in this paper is based on—in which visually impaired users employ the system to browse for audio books by pointing directly with their index fingers on the spine of the book as it stands in the shelf, drawing on similar interaction styles in other domains [3]. Recognizing which book it is the system very quickly provides the user with auditory information about the author, the title, and a summary of the contents of that particular book.

Third, before committing to development of a fully functional system we developed a Wizard-of-Oz prototype to be able to test and evaluate the concept with users early on in a real library environment. Intended users were thus allowed to experience using of the system, with a small amount of invested effort from our side. The results of these tests provided positive feedback from the intended users and a number of ideas for improvement.

Fourth, from these tests, we set out to develop a finalized prototype system in eight copies that were installed in Umeå City Library in March 2006 and is now in full operation.

3 Prototype Implementation Process

The finalized AudioIndex prototype was developed over the course of about a year. Compared to typical research prototypes, there were a number of considerations. First, as the system is to be used by the general public on a day-to-day basis, the system has to be robust and fail-safe both when it come to its physical design as well as to its software implementation. Second, as the system will be managed by librarians, with limited knowledge about Bluetooth, RFID, and wireless LAN

technologies, it also has to be easy to recharge, trouble-shoot, and reset. Third, as we were to make eight copies of the system, it had to be reasonable in terms of the cost of each device.

The part of the AudioIndex prototype that is utilized by the primary user, the library visitor, consists of three interconnected physical parts. First, the main device—worn around the neck using straps—contains a Dell Axim X30 PDA. The PDA is however fully enclosed in a custom-made plastic container, thus not being exposed to the user. Second, an earpiece provides the auditory information to the user. Volume is controlled using a small control attached to the cable. Third, a small pointing device is worn on the index finger. It consists of a custom-made plastic cover holding a battery and a dismantled MicroSensys iID PEN, connecting wirelessly using Bluetooth to the main device. The infrastructure of the system is largely hidden from the primary users, consisting of a SQL database, a wireless LAN, RFID tags, and a device that facilitates entry of information into the database.

Fig. 1. The AudioIndex prototype in use at Umeå City Library, Umeå, Sweden

3.1 Deployment Environment

The system was to be installed and used at the Umeå City Library, a public facility with around 2000 visitors per day housing an audio book section that contained about 3500 titles in DAISY format [2]. A DAISY book physically consists of a CD in a plastic case (20 x 171 x 144 mm). When delivered to the library, a DAISY book is fitted with stickers on the front and on the spine that provide information about the book in printed text and sometimes in Braille. It is also fitted with a barcode that is used for logistical purposes.

The entire public area of the library is served by a wireless LAN (WLAN). After the end of the project the test units were to remain in the library and be handled by the library staff, which would also have to maintain the database since the number of DAISY books grows by 20-30 volumes per week.

4 Hardware and Software Implementation

As the user points at the spine of a specific audio book, the RFID reader in the finger-mounted pointing device reads the identification number of the iCode-type RFID tag embedded in the book's spine and sends it to the PDA using Bluetooth. The PDA then sends a request over wireless LAN to a SQL server asking for all data for that particular book ID. After receiving this data, the PDA device employs speech synthesis software [5] to produce an audio stream that is fed to the user through the earpiece in the order of author, title, and summary. The entire system has been implemented in C# using Microsoft Visual Studio .NET 2003.

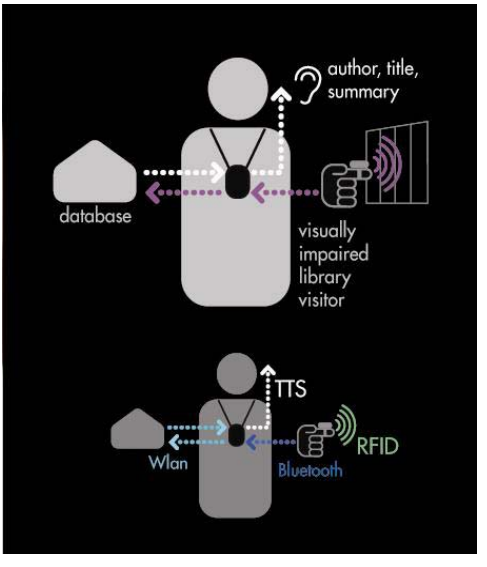


Fig. 2. The AudioIndex system overview

4.1 RFID Requirements

We wanted an RFID system that was small enough to fit the problem domain. The objects to be tagged were DAISY books as described above and the tag had to fit in the spine of the book (20 x 171 mm). For a visually impaired person the sense of touch takes on some of the tasks that the sense of vision performs for a fully sighted person. An important feature of the design of the reader is therefore that it can be worn on the index finger to allow users to use their fingertips as they would normally do. Therefore the reader needed to be fairly small. The spatial organization of the tagged objects; plastic cases with a 20 mm wide spine, side by side; set another requirement: that the reader would be able to discriminate the tag at which it is pointed from all other nearby tags. Some RFID readers are unable to read any tag if several tags are within the reading range simultaneously, other readers report all tags that are within the range. We obviously only wanted the tag that was pointed at to be read.

As far as possible, we also wanted to comply with any current or future standards for RFID use in Swedish libraries, and since we were going to deploy eight systems, the cost-per-device was also a limiting factor.

To deal with these requirements, we chose the Philips ICODE SLI protocol since ongoing discussions in Sweden regarding standardization of RFID in libraries point towards a solution heavily influenced by Danish standards, where ICODE SLI is the chosen RFID standard. We chose the MicroSensys iID ® PEN bt reader because this was the only reader available at the time that was small enough, wireless, and supported the ICODE SLI protocol. The tags chosen were I-code SCID PS from the German distributor Pavcard. The tags have 384 bits of free memory, a maximum reading range of 10 cm and support for anti-collision. What were more important to us were the dimensions of the tags: 25 x 9 x 0.3 mm, a low enough silhouette to make it possible to place them on the inside of the spines, thus making the whole solution more robust. Another important feature of these tags was that they were self adhesive.

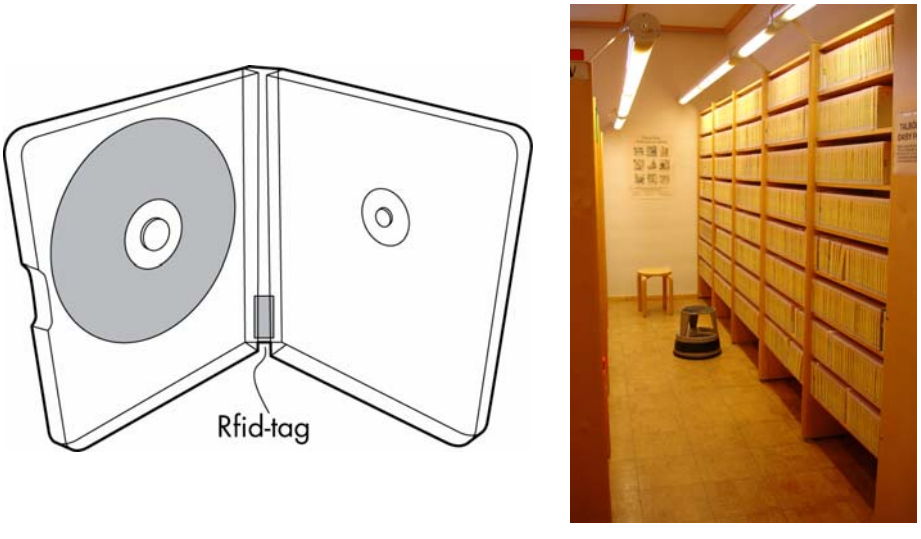


Fig. 3. The tagged object, a plastic case containing a CD (left); Spatial organization of tagged objects (right)

4.2 Distribution of Information

There were three options to choose from when deciding where to store the actual information content of the system. The first option was to store information about each audio book on the attached RFID tag. This option was quickly discarded since the information that would be presented to the user typically had a length of more than 400 characters per object (not counting overhead information necessary for subdividing data). In some cases the information exceeded 500 characters in length. 384 free bits would not nearly suffice unless we employed some compression algorithm, and with such a strategy we would still probably be left with hardly any space left for future additions to the data set.

The second option would be to deploy identical local databases on all PDAs. This would theoretically minimize the time to retrieve information from the database since the database resides on the same device that the retrieve requests originate from. However, this also creates a synchronization problem since the data set grows by about 20 to 30 audio books per week. The eight units would not be docked to a single computer since they needed to be distributed over the library premises and be handled by different librarians. It would have been possible to write code that would copy the most up-to-date database from a central server to each PDA at start-up time. We chose not to do this for a number of reasons. Neither the design team nor the test users experienced such a latency from the moment that a tag was read to the moment when the information was spoken that we felt it necessary to invest project resources to achieve a potentially lower latency. Throughout the project we have used user experience rather than absolute numbers as the meter by which to gauge the performance of our solutions.

The server side of the system needed to be deployed onto a server provided by the municipality for organizational reasons. This meant we had to comply with guidelines set by the IT-department of the municipality. These guidelines restricted our freedom in terms of software solutions and deciding not to write any server side code made it possible to steer clear of the thorn-field of writing fully compliant code.

The third and chosen option was to implement the database as a very simple SQL database containing only one table with seven columns. Whenever information needs to be retrieved from the database a request is sent via WLAN and the response is then returned.

4.3 Power Considerations

A typical visit to the audio book section does not last more than 30 minutes but we strove to maximize the time a unit could be used. The PDAs were fitted with a high capacity battery (1700 mAh compared to 950 mAh of the out-of-the-box battery) and by programming we turned off the screen of the PDA as soon as the PDA had established contact with the finger worn reader and the SQL database. We did not conduct an exhaustive test of battery capacity but after two hours of continuous use the battery meter indicated that it still held 50% of its capacity. Our tests with real users indicate that two hours of continuous use is enough for our purposes.

4.4 Creating the New Database

Creating and maintaining the database would be the responsibility of the library staff but we were responsible for finding a suitable hardware and software solution. Knowing that the number of audio books to be entered into the database was 3500 we realized that each second per book that we could save during the input process would be very useful for the library staff. Therefore, we spent a fair amount of time coming up with a solution that would automate as much of the work as possible. We built a custom reading unit that contains a barcode reader and an RFID reader placed in such a way that when an audio book is inserted, the barcode and the RFID code are read and used in a search-and-enter operation that only requires user involvement in special cases when something goes wrong. The typical scenario is that the operator

attaches an RFID tag to the inside of the book spine, inserts the book into the reading unit, waits a second or so until information about the book and a message about successful insertion into the database are displayed on the screen. When the book is removed, the information is also removed from the GUI. When dealing with several books, the library staff typically first attaches tags to a pile of books, then inserts them one after the other into the reading unit.

The key to getting hold of the desired information is the barcode, a subpart of which is a seven digit code identifier that we use to search an online database and from that retrieve information about author, title, summary, year of publication, etc.

4.5 Interaction

There are two roles, the library staff and the library visitor. When a library visitor wishes to use an AudioIndex unit, the librarian activates the finger worn reader by pressing a small button on its side, resets the PDA and then waits for the PDA to establish contact with both the reader and the database. When this contact has been established, the screen of the PDA is turned off and a red LED on the reader starts blinking. The librarian then inserts the PDA into its case which is worn like a necklace, and attaches the audio cable from the headphones. When the unit has been handed over to the visitor, he or she only has two interaction possibilities; point-and-listen or control the audio volume by turning a wheel on the necklace. When the visitor is done using the unit, the librarian must charge the batteries of both the PDA and the reader in order to ensure that the next visitor will be able to use a fully charged unit.

5 Discussion

An important, overall design goal of this project has been to hide interactional as well as computational complexity from the user. The only means by which the user interacts is actually by pointing at things in the physical library environment to get an audio description about the object, or, whenever the user wishes, immediately silence any current audio stream by pointing at the main device itself—all other screens, knobs, and buttons that make up the system are hidden from the user.

Due to the constraints that the system would have to fit into the informational infrastructure of the library and the local municipality, as well as into the daily routines of the library staff, we understood the importance of keeping the project grounded in the reality of the library and its staff. Many implementation decisions have been influenced by and directly related to the necessity of keeping the system realistic in its application context. This is not a bad thing however, since our aim in this project has not been to display our technical prowess but rather create a service that would benefit the library and its visitors today, without the need of a co-present research team facilitating its use. The challenges in such a project are different compared to those encountered in technology-driven explorative projects, but not less interesting. Some of the questions and challenges that we have dealt with in this project are discussed below:

Why create a new database instead of piggybacking on the existing database? The library already had a database containing most of the information that we needed to use during the project. Therefore, we explored the possibilities of adding two missing columns to the database; one containing RFID codes and one containing the book summaries. This proved futile for two reasons; it was not technically possible to extend the database and the copyright for the summaries did not allow that information to be added to library database. We were however allowed to use the summaries in a new project database separate from the library database.

Why strive for wireless operation? Wires tend to become rats' nests and thus become hard to handle even for visually enabled people. In our case, some of the librarians as well as the end users were visually impaired, so ease of use through wireless operation had high priority. In the end, we had to settle for headphones connected via wire to the PDA since we experienced severely impaired WLAN connectivity when using Bluetooth headphones. No detailed analysis of this phenomenon was made due to time limitations.

Why not build custom hardware ourselves? Time was a limiting factor, so instead of the to HCI design proverbial process of inventing a slightly different wheel, we opted for off-the-shelf hardware components, thus sacrificing optimal size in favor of time.

Additionally, to duplicate a prototype system tends to be quicker if the hardware consists largely of off-the-shelf products. Replacing faulty parts are likewise made significantly easier by using standard hardware—which is even more true if such an operation is to be performed by someone else than the researchers who built the system.

Why use WLAN to communicate with SQL server? The Umeå City Library already had a good WLAN infrastructure so it was natural to use that resource instead of building another kind of network, e.g. based on Bluetooth.

Why use a PDA? Other options could have been a mobile phone or a miniature computer like Oqo. At the time of the prototype project inception (early 2005) UMPC-grade computers had not been launched. Ideally, we would obviously have wanted to do away with the PDA altogether and include speech synthesis and WLAN capabilities in the finger worn device itself, but since such a product was not readily available and we did not have time or inclination to construct that product within the frames of this project, we wanted to get the smallest possible device that would suit our technical requirements within a reasonable budget. When decisions regarding hardware were also made at a time when there were no mobile phones available with WLAN, a speech synthesis that was accessible to third part developers, suitable Bluetooth profiles, and an IDE (integrated development environment) with a low learning threshold. We also had previous experience of developing PDA applications and knew that we did not really have time to spend learning a new IDE.

Why use speech synthesis rather than pre-recorded speech? Why was the speech synthesis done on the PDA? Pre-recorded speech would have given the users an improved experience both in terms of usability and in terms of pleasure. Speech

synthesis is generally still not as pleasant to listen to or as easily understood as human speech [cf. 5]. However, since there were no such audio files available, we would have had to create these ourselves and, subsequent to the installation, the library staff would have to spend considerable time creating such audio files. Also, such a system would in the long run be impractical to maintain, requiring new recordings to be made for every added item.

Speech synthesis software is executed on the PDA to avoid high latencies in the delivery of spoken information from the time of pointing. It was hard to estimate the load on the WLAN that was used for a number of other purposes and to minimize the load we opted for delivery of text from the database and conversion to speech on the PDA device rather than streaming audio which would induce a higher load on the limited resources provided by the WLAN.

Why use the synthesized voice 'Ingmar'? 'Ingmar' is the most widely used synthesized voice in Sweden so we expected users, who often use speech synthesis on a daily basis, to feel familiar with Ingmar. This was also a sentiment that was voiced by test users throughout the project. We evaluated several Swedish synthesized voices but found that the one that was most accepted was Ingmar. Subsequently to the installation of AudioIndex in March 2006, another, more life-like voice has become possible to use on PDA; Erik.

Why not make it possible to dynamically alter the speed of the speech synthesis? Many experienced users of speech synthesis technology become so adept at interpreting the voices that it is possible to increase the speed of the speech by as much as 50% in some cases. Once these users have gotten used to a higher speed, the normal speech synthesis speed seems excruciatingly slow.

In most applications containing speech synthesis it is therefore wise to provide a mechanism by which the user can alter the speed as he/she sees fit. As discussed above, we however wanted to keep the interaction simple so we did not really want to add anything to what we already had, i.e. the ability to point at objects to receive information about those objects and the possibility to control audio volume through a small control wheel on the necklace holding the PDA. Therefore we decided to not let the user alter the speech synthesis speed dynamically but instead be able to choose a higher speed when he/she borrows the unit from the library. This was implemented by us actually hard-coding a slightly higher speech synthesis speed (20% higher than normal) into some of the units. While possibly not an ideal solution, at least users are able to tell the library staff whether they want a high or normal speed unit. An alternative design could have been to provide a 'configuration table' with Braille text next to the library desk, allowing users to customize their device by pointing.

6 Conclusions

We have presented the AudioIndex prototype, part of a larger effort in making public libraries more accessible to visually impaired. It allows users to browse and search for audio books without the need for library staff guidance. AudioIndex is a light-weight mobile system based on a combination of using RFID, PDA, and wireless

technologies, and allow users to point at objects (typically audio books and bookshelves) and get audio feedback (typically in the form of synthesized speech).

This project challenged the design team in several ways. First, the parts of the system exposed to the users needed to be robust and easy to operate since the target group was varied in terms of age, background and most other characteristics apart from being visually impaired in some way, though such impairments can also be quite different and hence be more or less impairing. This requirement was fulfilled by disallowing any other interaction with the system apart from pointing at objects (typically audio books and bookshelves) and listening at the associated information. Hiding electronic parts of the system inside audio book covers and injection molded plastic casings also minimized the risk of unintentional damages to more delicate parts of the system. Second, the system had to fit into the present library environment both physically and in terms of routines and informational architecture. RFID protocols were chosen so as to harmonize with emerging Swedish library standards for RFID and the construction and maintenance of the system infrastructure utilized existing solutions wherever possible or, if not possible, was designed in such a way that it would minimize interference with other routines. Third, the cost of development was an issue, especially since we were to deploy eight copies of the end user devices. We chose to work with hardware and software modules that required only slight tweaking to provide the desired functionality and focused our own efforts on creating the software, digital glue if you will, that tied all the pieces together to form a whole system.

Acknowledgements. Thanks to Louise Maniette and Anna Karin Bergkvist.

References

1. Bradley, N.A., Dunlop, M.D.: An Experimental Investigation into Wayfinding Directions for Visually Impaired People, Personal and Ubiq. Comp. vol. 9(6) (2005)
2. DAISY: Digital Accessible Information System (DAISY) Consortium. From the World Wide Web (Retrieved February 13, 2007) http://www.daisy.org/
3. Fallman, D.: Wear, Point, and Tilt. In: Proc. of DIS2002 (London, UK, June 25-28), pp. 25–28. ACM Press, New York (2002)
4. Fallman, D., Lindbergh, K.: Inclusive design by taking special measures — making libraries accessible for all. In: Proceedings of Include 2007 (Royal College of Arts, London, UK, April 2-4) (2007)
5. Pitt, I.J., Edwards, A.D.N.: Improving the usability of speech-based interfaces for blind users. In: ACM Conf. on Assistive Technologies (Vancouver, Canada), ACM Press, New York (1996)

Performance Evaluation of Voice Interaction as a Universal Web Interface

Yu-Ming Fei, Chiuhsiang Joe Lin*, Min-Ting Chen, and Chih-Cheng Chiang

Department of Industrial Engineering, Chung Yuan Christian University
200, Chung Pei Rd., Chung Li, Taiwan 32023, R.O.C.
* hsiang@cycu.edu.tw

Abstract. A speech interface sometimes provides easy access to WWW information since it makes browser potentially more friendly and powerful. This study first made a comparison between a voice and a keyboard input task. It then compared the voice interface with the keyboard, mouse, and combination of keyboard and mouse with three WWW browsing tasks. The experimental results showed that the voice interface performed the input task faster than the keyboard. The voice interface reached an input accuracy of 92% after sufficient learning. In the WWW browsing tasks, the voice performed poorly in terms of completion time, compared to the keyboard, mouse, or both. Nevertheless, the voice interface can be used to complete the three assigned tasks without problems. The study showed that voice interaction can be used as a universal web interface, especially for those who are unable to use the mouse and keyboard due to disabilities.

Keywords: Speech interface, Voice interface, IBM ViaVoice, Browse WWW.

1 Introduction

It is very common to find information through internet service. Since internet can supply huge database and can supply a lot of information at any time. Normally, people are familiar with the keyboard and mouse as the major interfaces, but for people such as old people or people with disability, it becomes difficult to access the internet service. For people who cannot use the keyboard and mouse, it is important to develop other methods for them to access the internet.

Software that can convert spoken words into written text (hereafter referred as SWIWT) has been available since the early 1980s. Early continuous speech systems were first introduced in 1994, with the latest commercially available editions having a claimed accuracy of up to 98% of speech recognition at natural speech rates [1]. The IBM software was found to have the lowest mean error rate for vocabulary recognition (7.0 to 9.1 percent) followed by the L&H software (13.4 to 15.1 percent) and then Dragon software (14.1 to 15.2 percent) [2]. In addition, the IBM software was found to have better performances than both the Dragon and the L&H software in the recognition of general English vocabulary and medical abbreviations [8].

* Corresponding author.

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 279–288, 2007.
© Springer-Verlag Berlin Heidelberg 2007

In spoken words into written text, many researches have been doing in this field to explore the reliability measurement. Wang, Chou, and Chen [6], pointed out that the interface currently provides speak-able commands, bookmarks and links. However there is no general literature pointing out what kind of tasks can be implemented when browsing the web. This study attempts to use an experiment design to verify the task performances on the web using the IBM Via Voice software.

This research first makes a comparison on the difference between SWIWT and keyboard inputs. According to the instruction of the software, SWIWT has its own learning ability. The experiment will only begin after system self-learning and after more voice modules have established. The research then carries out an experiment to view web interface and complete certain tasks by using different input interfaces which include mouse, keyboard, and speech.

2 Method

A two-phase study was designed to explore the performances between speech interface, keyboard, and mouse. The first study was to compare "spoken words into written text" (hereafter referred as SWIWT) with keyboard input. The second study was to compare the performances between speech interface, keyboard, and mouse in browsing the web pages under this experiment. There are 3 male graduates subjects participating in this experiment who study in Chung Yuan Christian University. They had a comparable level of computer literacy and web expertise. The following hardware and software were used: a desktop 3.0-GHz Intel Pentium 4 processor with 256 MB of RAM, noise-canceling headset microphone, and IBM ViaVoice Pro version 10.

2.1 Experiment 1

The main factor was two input text interfaces- SWIWT and keyboard. Each of these three participants had to test each input text interfaces for three times. The sequence of the experiment was nested under each input text interface, and the participants were designed to be randomized block. This study randomly selected an article (398 words) from TQC (Techficiency Quotient Certification) article database for all input tests. The main objective of this study is to compare the difference between SWIWT and keyboard input, and to explore the effects of testing between them.

Therefore the model for the nested design is

$$Y_{ijk} = \mu + \tau_i + \alpha_{j(i)} + \beta_k + \varepsilon_{ijkl} \begin{Bmatrix} i = 1,2 \\ j = 1,2,3 \\ k = 1,2,3 \end{Bmatrix} \tag{1}$$

$$\text{Input text interface: SWIWT and keyboard } \quad \tau_i \rightarrow i = 1,2 \tag{2}$$

$$\text{Sequence} \quad \alpha_{j(i)} \rightarrow j = 1,2,3 \tag{3}$$

Fig. 1. Figure 1: The participants used SWIWT and keyboard to input texts

$$\text{Participants} \quad \beta_k \rightarrow k = 1,2,3 \tag{4}$$

When using the SWIWT interface, it was different in the speech model training each time. In the first time, every participant must create a base of personal speech model, which took a long time. Before the second SWIWT input task, each participant had to spend about 35 minutes to read an article from IBM ViaVoice system to enhance the established personal speech model. Similarly, before the last SWIWT input task, another 20 minutes or so were used by each subject to enhance the personal speech model again.

This experiment recorded the completion time and the error words of the input tasks. The study also observed the effect of self-learning with the IBM ViaVoice system.

2.2 Experiment 2

This research then started the second experiment on web control. There were three tasks assigned in the experiment, and they were all universal access activities when browsing the web. The second experiment included two factors. The first factor is the tasks in browsing the web pages, containing the following three tasks: (1)Task one requires the subject to find out designated web and click on it from My Favorite. (2)Task two must trigger the required action on the web by moving the cursor. (3)The objective of the third task is to input data into appointed location on the web.

The second factor is the interface used to brow the web pages. There are four interfaces in this factor, 1) use keyboard and mouse, 2) just use keyboard, 3) use mouse only, 4) use speech interface. The three participants were designed to be the randomized block. In this experiment the time spent for each task was observed and recorded from the start to the end.

The statistical model for this design is

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \gamma_k + \varepsilon_{ijk} \begin{Bmatrix} i = 1,2,3 \\ j = 1,2,3,4 \\ k = 1,2,3 \end{Bmatrix} \tag{5}$$

$$\text{Tasks:} \quad \alpha_i \rightarrow i = 1,2,3 \tag{6}$$

$$\text{Interfaces:} \quad \beta_j \rightarrow j = 1,2,3,4 \tag{7}$$

$$\text{Participants:} \quad \gamma_k \rightarrow k = 1,2,3 \tag{8}$$

2.3 Task and Instructions

The participants were told prior to experiments as to the nature of the study that we were interested in determining whether the tasks are better in terms of time to complete them. However, they were not told what types of results were expected. Participants have one hour practicing before carrying out the experiment, and they are requested to be familiar with all the voice commands.

Task One. The objective of task one is to find out designated web and click on it from My Favorite. The procedure is as follows.

1. Open browser and use my favorite to go to the page, e.g. "健康醫 " ("Health Web").
2. Find out the "News" of the health link and go to the fist news link. Visit the page to the bottom, and then use the "Back" to go to "News".
3. Find out the first news on "Health of residence", and visit the page to the bottom and then say the last sentence.

Task Two. The objective of task two is to trigger the action on the web by moving the cursor, as described in the following.

1. Use my favorite to go to the page, e.g. " 界" ("The Flower World"), and then go into the main page from the door picture in the center of web page.
2. Slide the cursor to move about the four flower pictures, and make them to change to another picture. After flower pictures changed, keep sliding the cursor to point to the " 之 " ("Symbolization of Flowers") button underneath the page.
3. Find out the symbolization for tomato in the "Symbolization of Flowers" page, and then move the cursor to point to the " 馬看 " ("Poems of Flowers") button underneath the page.
4. Browse through the "Poems of Flowers" page to the bottom and then say the last sentence.

Task Three. The objective of task three is to input data into appointed location on the web.

1. Use my favorite to open the page, e.g. " 見 " ("The suggestion"), and then keep the cursor on the form fill-in.
2. Type "王 " ("Wang Xiao-Ming") for name information, type "54321" for password, then select "Female" for sex, and type " 喜 " ("A beautiful web design and I like it very much.").
3. Press the "Send" button when finished with the typing task.

3 Results

This section presents and discusses the results of the two experiments. For each experiment, the data were assessed as to the fit of the normality assumption, the independence assumption, and the equality of variance. Then, summary measures and tests for main effects and interactions are presented on completion time of tasks across levels of experimental factors. The data analysis was performed on a Pentium 4 PC using Minitab release 14.

3.1 Experiment 1

The ANOVA summary table for data on completion time (Table 1) indicates a significant main effect for input text interface (F = 57.25, P = .000). However, there was no significant main effect for sequence (F = .48, P = .749). The results suggest that SWIWT (M=144 seconds, SD=7.06 seconds) used less completion time than keyboard to input texts (M=688 seconds, SD=238.3 seconds). The SWIWT showed a more consistent response in task speed among the participants as reflected by the smaller standard deviation.

The ANOVA summary table for error words (Table 2) indicates a significant main effect for input text interface (F = 19.77, P = .001). However, there was no significant main effect for sequence (F = 2.26, P = .135).

A good way to further analyze the data is the analysis of means (ANOM). This test not only answers the question of whether or not there are any differences among

Table 1. Analysis of variance for completion time, using adjusted SS for tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Input text interface	1	1331712	1331712	1331712	57.25	0.000
Sequence	4	44781	44781	11195	0.48	0.749
Participant	2	180652	180652	90326	3.88	0.056
Error	10	232609	232609	23261		
Total	17	1789754				

S = 152.515 R-Sq = 87.00% R-Sq(adj) = 77.91%.

Table 2. Analysis of variance for error words, using adjusted SS for tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Input text interface	1	5477.6	547736	5477.6	19.77	0.001
Sequence	4	2505.6	2505.6	626.4	2.26	0.135
Participant	2	398.8	398.8	199.4	0.72	0.510
Error	10	2770.6	2770.6	277.1		
Total	17	11152.4				

S = 16.6450 R-Sq = 75.16% R-Sq(adj) = 57.77%.

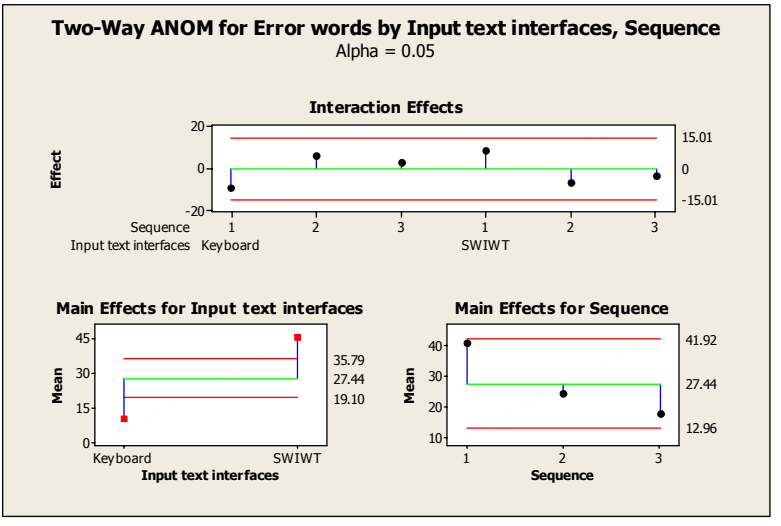


Fig. 2. Two-way ANOM for error words by Input text interface, Sequence

the factor levels, but also tells us which levels are better and which are worse when there are differences [3]. For the sake of providing visual picture of the distinction between factor levels, consider the graphic representation in Figure 2. The number of error words between keyboard and SWIWT were strikingly different, with the error words of keyboard fewer than SWIWT. Since the self-learning ability of SWIWT was most concerned in this study, we observed a decreasing trend in the number of error words for sequence, as can be seen in Figure 2. Since in the original model of ANOVA, sequence was nested under interface, no interaction between input text interface and sequence was assumed. If we only looked at the sequence effect under the SWIWT interface, Table 3 further shows the mean error rate for sequence. Clearly, the data show that the time spent in enhancing the personal speech model had a positive effect on participants performance to reduce error rate. The result indicate that the self-learning feature of the IBM ViaVoice system could improve accuracy rate to above 92%.

Table 3. The mean error rate of sequence in SWIWT

	First time	Second time	Third time
Error rate	16.83%	8.83%	8.00%

3.2 Experiment 2

In this experiment we wish to investigate the effect of single and multimodal interfaces, which combine two modes from the keyboard, mouse, or speech for browsing web page tasks. But in the study, task two (moving the cursor) could not be performed by using the keyboard only. Therefore, the analysis was divided into two

parts. The first analysis leaves out the data on task two, 'moving the cursor'. The second analysis leaves out the data on the interface, 'just use keyboard'.

Results by Excluding the Task of Moving Cursor. The results on the performance of completion time are presented in Table 4. Analysis of variance (ANOVA) showed a significant effect of task, $F(1,14)=158.89$, $p < .001$; a significant effect of interface, $F(3,14) = 145.16$, $p < .001$; and an interaction between task and interface, $F(3,14) = 29.81$, $p < .001$. A more detailed understanding of the relationship between these effects can be gained from Figure 3. Using speech interface to browse through the web was the most difficult and wasted a lot of time than the other interfaces.

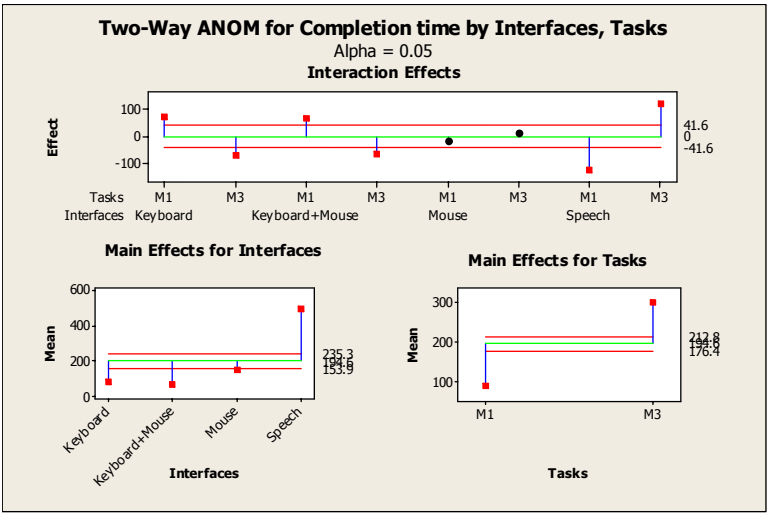


Fig. 3. Two-way ANOM for completion time by interface and task (M1: Task one, M2: Task two, M3: Task three.)

Table 4. Analysis of variance for completion time, using adjusted SS for tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Task	1	267126	267126	267126	158.89	0.000
Interface	3	732147	732147	244049	145.16	0.000
Task*Interface	3	150378	150378	50126	29.81	0.000
Participant	2	4666	4666	2333	1.39	0.282
Error	14	23537	23537	1681		
Total	23	1177854				

S = 41.0028 R-Sq = 98.00% R-Sq(adj) = 96.72%.

Results by Excluding the Only Keyboard Interface. As shown in Table 5, analysis of variance (ANOVA) showed a significant effect of task, $F (2, 16) =368.00$, $p < .001$; a significant effect of interface, $F (2, 16) = 113.05$, $p < .001$; and a significnat

interaction between task and interface, F (4, 16) = 22.37, p < .001. A more detailed understanding of the relationship between these effects can be gained from Figure 4. The speech interface to browse through the web was the most difficult and wasted a lot of time than the other interfaces.

Table 5. Analysis of variance for completion time, using adjusted SS for tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Task	2	992544	992544	496272	368.00	0.000
Interface	2	304900	304900	152450	113.05	0.000
Task*Interface	4	120653	120653	30163	22.37	0.000
Participant	2	8348	8348	4174	3.09	0.073
Error	16	21577	21577	1349		
Total	26	1448022				

S = 36.7229 R-Sq = 98.51% R-Sq(adj) = 97.58%.

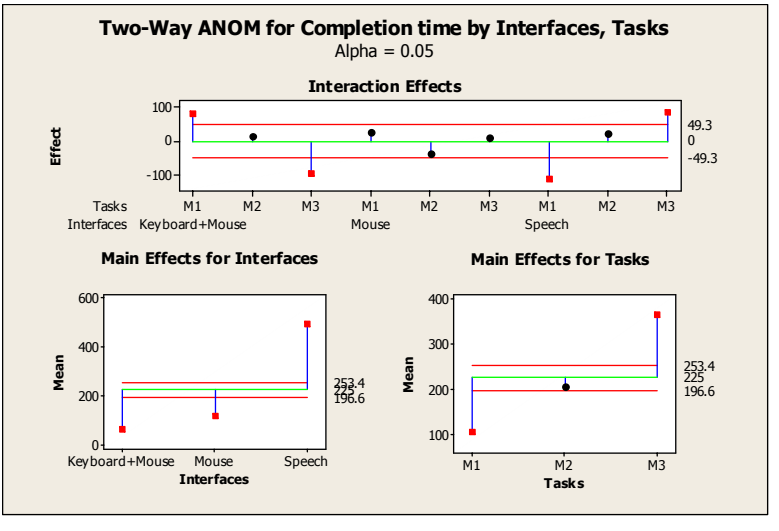


Fig. 4. Two-way ANOM for completion time by interface and task (M1: Task one, M2: Task two, M3: Task three.)

4 Discussion and Conclusions

From the results of the experiment, this study has found out that apparently less time was spent on the text input when using the voice interface in comparison with typing (non-voice interface). In aspect of error rates, although there were more errors on the text recognition system when initially started using the voice interface, incorrect texts were obviously reduced after voice training was given. According to the result of the experiment, voice interface did have the learning ability and its mean accurate rate

can reach up to 92% especially in later stage of the voice interface. Even thought error rate of voice interface is higher than typing, the speed of text input is quicker, and the voice interface is still very practical.

The error rate was decreased through system learning by enhancing personal speech models. In 1999, a study of ASR (Automatic speech recognition) performance on broadcast news showed wide variations, from a maximum of 88% words correctly recognized to a minimum of 35%, with a mean of 67% [7]. A recently research of ASR performance in the word error rate of the overall integrated system could reduce to 9.26% to 9.61% [4]. Up to this point, these results show ASR still advancement. An additional result was 165.8 words per minute in this study. It may be of interest for future research on how to increase speech speed and accuracy.

To summarize the salient features of the analysis, several findings are of interest. This study revealed that improving method of operation appears to be an important variable in browser WWW. The speech interface could cost a lot time even though it can do most of job.

For the assigned tasks in this study, it is more difficult to control cursor and to browse web when using the voice interface. However, it does provide an advantage for disabled users. The performance evaluation of voice interaction as a universal web interface can help those who are unable to use the mouse and keyboard to achieve the same result as using the mouse and keyboard.

For WWW applications, the continuous reaction time must be immediate and smooth. Participants were anxious when they used the IBM ViaVoice to shift the cursor. Often users like to find some interesting part to browse around a web page. Using speech to browse a text link had fast response time, but moving the cursor to point to a link took a long response time, because there was a need to keep using the voice to shift the cursor even though some commands could be retrieved fast. Shneiderman, and Plaisant, [5] pointed out that an appropriate response time on typing, cursor motion, or mouse selection was around 50-150 milliseconds. It seems that there is still a big gap for the voice interface to be realistically used in the future.

Another point is that a popular web needs to be designed for universal web Interface. The study found that there was no way to move the cursor on a web page if usuing the keyboard only. Voguish web design uses a lot of skills and techniques, but usually does not consider what possible universal interfaces will be used. For popular web pages, designing should be suitable for all multimodal interfaces (keyboard, trackball, speech, pen, touch…). Further WWW design issues should be pursued toward a handicap-friendly place on the Internet.

Although the sample in the study was small, the following recommendations could serve as some general principles for researchers who would like to experiment with ASR in similar contexts. The results of the present study may be concluded by pointing out that ASR could be applied to Web Interface but still needs some pratical improvements.

References

1. Al-Aynati, M.M., Chorneyko, K.A.: Comparison of Voice-Automated Transcription and Human Transcription in Generating Pathology Reports. Archives of Pathology and Laboratory Medicine 127(6), 721–725 (2003)

2. Devine, E.G., Gaehde, S.A., Curtis, A.C.: Comparative Evaluation of Three Continuous Speech Recognition Software Packages in the Generation of Medical Reports. Journal of the American Medical Informatics Association 7, 462–468 (2000)
3. Nelson, P.R., Coffin, M., Copeland, K.A.F.: The Analysis of Means. Introductory Statistics for Engineering Experimentation, pp. 250–260. Elsevier Academic Press, North-Holland, Amsterdam (2003)
4. Schwenk, H.: Continuous space language models. Computer Speech and Language. In Press, Corrected Proof, Available online October 9, 2006 (2006)
5. Shneiderman, B., Plaisant, C.: Designing The User Interface: Strategies For Effective Human-Computer Interaction, 4th edn. p. 473. Addison Wesley, USA (2005)
6. Wang, H.M., Chou, Y.H., Chen, B.: Browsing The Chinese Web Pages Using Mandrin Speech. International Journal of Computer Processing of Oriental Languages 13(1), 33–50 (2000)
7. Whittaker, S., Hirschberg, J., Choi, J., Hindle, D., Pereira, F., Singhal, A., SCAN,: Designing and evaluating user interfaces to support retrieval from speech archives. In: Proc. ACM SIGIR '99 (1999)
8. Zhang, J., Zhao, J., Bai, S., Huang, Z.: Applying Speech Interface to Mahjong game. Multimedia Modelling Conference. In: Proceedings. 10th International, pp. 86–92 (2004)

Fisheye Views of Java Source Code: An Updated LOD Algorithm

J. Louise Finlayson, Chris Mellish, and Judith Masthoff

University of Aberdeen
{lfinlays,cmellish,jmasthof}@csd.abdn.ac.uk

Abstract. One very important aspect of computer programming is reading source code. Whilst this is a relatively simple process for sighted programmers, for blind computer programmers this presents a significant problem. Navigating through and comprehending often thousands of lines of code can be time consuming and difficult. Current development environments have many features that aid the reading of source code for sighted users, however, most of these features are visual in nature and are not translated well into audio by general screen-reader applications. Research has suggested that the use of fisheye views could aid navigation and enhance performance in program comprehension activities for blind programmers. This paper reports the results of a study into creating a better fisheye view of Java source code, by improving the method used to determine each line's 'global importance' or 'Level of Detail' (LOD). The traditional LOD determination method uses only the indentation level of a line to calculate its overall importance. This paper describes the results of the study, and suggests some of the issues which may need to be considered in developing an improved LOD calculation for programming source code.

1 Introduction

Visual systems for displaying and editing program source code use a number of different methods to try to aid program comprehension. Many of these methods act by suppressing unwanted information and highlighting the essential. One such solution, the corollary of a 'zoom' lens allows the user to focus on one specific area in detail whilst ignoring the rest of the information. This approach is not ideal, as noted by George Furnas [5]:

"In such a zoom system local and global information are never available at once. Integration of such information must take place in human memory"

He also discusses the problems involved in a 'two-window' set-up where one window shows the low level minutiae, whilst the other shows less detailed information:

"…this approach has problems, too, basically in understanding the correspondences between the two views. Where does the small view fit in the big picture and what visual features correspond?"

Furnas proposed a novel solution to this problem using a technique known as the 'fisheye' view. A fisheye view [5, 6] is a method used to tailor the level of detail

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 289–298, 2007.
© Springer-Verlag Berlin Heidelberg 2007

shown at, and around, the user's focal point. This is achieved by calculating the user's likely interest in each line, given their current position, and then displaying or suppressing it accordingly. In this way, the fisheye view of a program is constantly redefined as the user changes their focal point. This method allows the user to view detailed information at their points of interest whilst also displaying the surrounding contextual information.

There is a substantial amount of evidence that supports the assertion that this type of view is more effective than any of the alternative methods mentioned previously. Schaffer et al [7] showed that, compared to the full-zoom method, the fisheye view allowed users to navigate more quickly and retain information about previously navigated spaces more easily. There have also been previous studies which show an increased effectiveness of the fisheye view when compared to other methods such as 'pan-and-zoom' [2, 4]. There has been much interest in the fisheye concept since its creation and at least two program editing environments 'emacs' [8] and 'Jaba' [1] incorporate fisheye lens view features to aid the browsing of program source code.

Our work investigates the use of auditory fisheye overviews to aid navigation and enhance performance in a non-visual program comprehension task. The hypothesis is that providing a glance at the source code in this way will improve the readers' understanding of the program.

2 The Standard Fisheye Algorithm

A fisheye view can be thought of as a 'Degree of Interest' (DOI) filter. The DOI value represents how interested a user is in seeing any particular line, given their current position. The fisheye view defines the DOI of various parts of the structure by taking into account the current focal point, the global importance of each of the lines (LOD) and the structure's proximity to the current focus (see figure 1). In terms of Java code:

- The focal point is the line on which the cursor is currently located.
- The code is treated as a hierarchy, and the distance from the focus (D) is calculated as the path distance between nodes rather than the absolute distance between lines (shown in figure 2).
- The LOD of each line is calculated as a function of its level of indentation, with less-indented lines deemed to be more important than nested structures.

Using these three properties, the DOI function at point x, given the current focus '.' can be defined as:

$$\mathbf{DOI(x|.) = (LOD(x) - D(.,x))} \quad (1)$$

That is, interest increases with global importance, and decreases with distance from the focus. This is the original algorithm developed by Furnas, and used by both Jaba and emacs to generate their fisheye views. Typically, to create the fisheye view, the line x is displayed only if its DOI is above some predetermined threshold. Figure 1, below shows an example of the DOI function calculation for an excerpt of nonsense source code. For each line, LOD is calculated as (indentation level +1) * -1.

For example, line number 11 (at two levels of indentation) is assigned an LOD of (2+1)*-1 = -3.

	Source Code	D(.,x)	LOD(x)	DOI(xl.)
1	XXXXX	3	-1	-4
2	XXXXX	4	-2	-6
3	XXXXX	2	-2	-4
4	XXXXX	1	-3	-4
5	XXXXXX	2	-4	-6
6	XXXXXX	2	-4	-6
7	FOCUS '.'	0	-4	-4
8	XXXXXX	1	-5	-6
9	X	0	-4	-4
10	XXXXXXX	2	-4	-6
11	X	1	-3	-4
12	X	2	-2	-4
13	XXXXXXX	4	-2	-6
14	X	3	-1	-4

Fig. 1. Example showing DOI function calculation for an excerpt of nonsense code

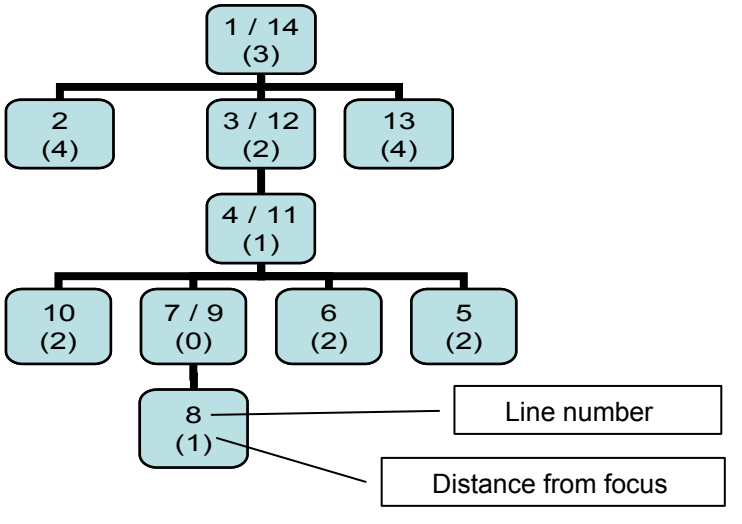


Fig. 2. Path distance calculation for each line of nonsense code (from figure 1)

In our own work, this algorithm was used in an investigation into the use of an auditory fisheye view to aid orientation and enhance performance in a non-visual program comprehension task. The investigation compared two prototype code-reading environments - one incorporating a 'flat' overview of the source code and the other

utilizing a 'fisheye' overview. In all other aspects, the two environments were identical. To create the 'flat' overviews each line was simply represented as its basic construct. For example, the following lines...

```
while (Thread.currentThread() == loopThread){
for (int i = 1; i <= 10; i++){
```

were summarized as a 'WHILE', and a 'FOR' respectively.

To create the 'fisheye' overviews, three levels of detail were represented using two DOI threshold values: A line with a DOI value above the higher threshold was spoken in full, a line with a DOI of less than the lower threshold was represented as a simple tone. Any line with a DOI falling between these two values was rendered as its basic construct (identical to the information displayed in the 'flat' view representation). Both systems used the JAWS for Windows screen-reader software to produce the auditory output.

To evaluate each interface the subjects were required to use one of the systems to navigate through and answer questions on a sample Java program. The time taken to answer each question was recorded, as was the number of correct answers, a log of the user's keystrokes and any comments made. After all the questions had been presented, the subject was asked to complete a NASA TLX workload evaluation to determine how difficult they judged the task to be, and a simple questionnaire about their impressions of the system. This procedure was then repeated with the alternative system, and finally, the subjects were asked to complete a third questionnaire comparing the two systems and indicating their preferences.

The results of this investigation showed a slight trend in favour of the fisheye system (63% of subjects preferred it to the 'flat' overview system and it was judged to be less demanding by the TLX evaluation findings). However, both of these results failed to reach statistical significance. The main finding of the investigation arose from the subjects' comments which revealed that the majority of subjects (88%) felt that the lines that were displayed or suppressed in the fisheye view were not always appropriate, and that important lines were sometimes suppressed whilst trivial lines were displayed in full. This problem is masked (yet not resolved) in some other systems by the heavy use of manual overrides (which allow the user to always suppress or display particular regions regardless of their calculated DOI).

Whilst there is undoubtedly a need for the user to have this manual control, it is essential that the system minimize the amount of work the user has to perform in order to use it, and that the default fisheye view be as appropriate as possible. The large number of negative comments on the information tailoring provided by the default fisheye calculation may indicate that the standard algorithm used is not the most effective in this situation.

The part of the algorithm most likely in need of review is the calculation of each line's LOD. The notion that the importance of a given line can be derived solely from its level of indentation seems to be simplistic at best and does not take into account any information about the type of construct or the content of the line. The study presented in this paper aimed to determine whether there is a more accurate way to assess the global importance of each line than by using the standard algorithm.

It was deemed important that any algorithm developed should only involve information about the source code that is readily available and does not involve an

understanding about the workings of the program. Having a deep understanding about the specifics of the program, and the purpose of every line would undoubtedly provide a very accurate measure of LOD but is an unfeasible requirement for a general algorithm. For this reason, the study concentrated on factors providing structural information such as construct type and indentation. This type of information is easily available to the system and is often already used by programming environments to produce syntax highlighting and other features.

This paper describes the results of the study, and suggests some of the issues which need to be considered in developing an improved LOD calculation for programming source code.

3 Methodology

3.1 Participants

Seventeen subjects took part in the study, all with some experience with the Java programming language. As the experiment did not involve writing code it was not essential that they be expert Java programmers, however, the participants had to be at least familiar with the different types of construct and Java syntax.

3.2 Procedure

Subjects were asked to look at an excerpt of Java code and to consider the overall importance of each line. They then rated each line's importance in contributing to an overall understanding of the program, on a scale of 1 (very unimportant) to 7 (essential). Each subject performed this task five times, with different samples of code. The average degree of importance given to each line, as determined by the subjects, was judged to be its genuine LOD. This LOD value could then be compared to the results of various LOD calculations to find the method that most closely correlated to it.

4 Results

The subjects' Java experience varied from 'use it occasionally' to 'expert'. No significant differences were observed between the answers given by subjects with different levels of experience.

No significant correlation was found between the assessed global importance of a line and its level of indentation ($r = -0.64$, $p>0.05$), see table 1 for the results. This was confirmed by ANOVA, which revealed no significant differences between any of the different levels of indentation ($F(5, 136) = 0.46$, $p>0.5$).

The next factor to be investigated was that of construct type. Earlier work on how best to provide an auditory glance at program source code [3] had revealed that Java could be considered as consisting of six general types of construct, shown below in figure 3.

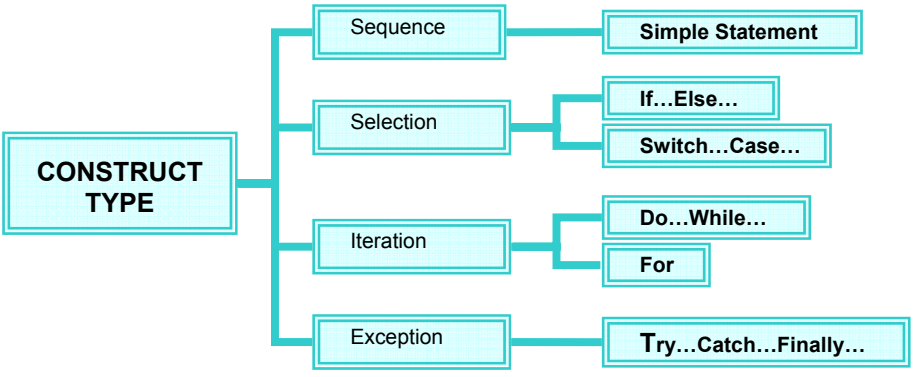


Fig. 3. Original basic Java Construct types

For the purposes of the study, this notion was expanded upon, and three further general categories were identified. These were: Class, Method and End. The Class and Method types include any lines that define either a class or a method respectively, while the End types represent the end of a nested construct, i.e. '}'. This created nine general Java construct types, as shown below:

Type 1 - Class	Type 4 - If...Else...	Type 7 - Try...Catch...Finally
Type 2 - Method	Type 5 - For	Type 8 - Switch...Case...
Type 3 - Simple Statement	Type 6 - Do...While...	Type 9 - End

Of these nine basic types, only seven were represented in the experiment (types 7 and 8 were not included). The mean, assessed Global Importance associated with the remaining seven construct types is shown in table 2. The construct types of the lines of code were compared to their calculated overall importance to determine if any relationship could be seen. The results of the ANOVA revealed that there were significant differences between the groups as a whole ($F(5, 136) = 103.51$, $p<0.0001$).

Further analysis using a multiple range test to determine the source of this variation showed that Type 3 ('simple statements') and Type 9 ('ends') constructs varied significantly from each other, and from the other construct types - with much lower assessed LODs.

With 'simple statement' and 'end' constructs varying significantly from all other types, we looked once more at indentation as a measure of LOD, this time excluding 'end' and 'simple statement' types from the calculation, see table 3 for the results. This time, only five different levels of indentation could be examined as the highest level of indentation can only be filled by 'simple statement' types and was therefore not included in this calculation.

It was shown that there is now a clear, linear correlation between the assessed global importance of a line and its level of indentation ($r = -0.92$, $p<0.01$). This was supported by ANOVA, which revealed there were significant differences between the five groups ($F(4, 36) = 13.77$, $p<0.0001$).

The indentation of 'simple statement' types and 'end' types as a measure of LOD was investigated. In both cases, there were no significant differences observed

between the groups ($p>0.05$), and no semblance of a correlation between a line's indentation level and its genuine LOD.

Type 3 'simple statement' constructs can be further classified according to their specific sub-type. The mean LOD associated with each statement sub-type identified in the experiment can be seen in table 4, below.

These sub-types of statement were compared to their judged level of importance to determine if any relationship could be seen. The results revealed that there were significant differences between the groups as a whole ($F(3, 66) = 13.11$, $p<0.0001$).

Further analysis using a multiple range test to determine the source of this variation showed that all four of these sub-types varied significantly from each other ($p<0.05$).

Table 1. Indentation level as indicator of global importance

	Global Importance	
Indentation Level	*Mean*	*St Dev*
0	4.74	2.06
1	4.51	1.36
2	4.24	0.96
3	4.48	0.92
4	4.28	0.87
5	4.40	0.00

Table 2. Construct type as an indicator of global importance

	Global Importance	
Construct Type	*Mean*	*St Dev*
Type 1 - Class	6.70	0.14
Type 2 - Method	6.03	0.56
Type 3 - Simple Statement	4.25	0.60
Type 4 - If…Else…	5.81	0.53
Type 5 - For	5.12	0.18
Type 6 - Do…While…	5.93	0.23
Type 9 - End	2.86	0.28

Table 3. Indentation level as an indicator of global importance excluding 'end' and 'simple statement' types

	Global Importance	
Indentation Level	*Mean*	*St Dev*
0	6.68	0.11
1	5.89	0.28
2	5.70	0.54
3	5.67	0.58
4	4.40	0.00

Table 4. Sub-class as an indicator of global importance for 'statements'

	Global Importance	
Statement Sub-type	*Mean*	*St Dev*
Method Call	3.72	0.27
Declaration/Assignment	4.73	0.47
Return Statement	4.20	0.28
Other	4.01	0.53

5 Discussion

If the standard fisheye view algorithm were to be substantiated, it would be expected that the indentation of each line would predict the level of importance given to it by the subjects. This did not occur, however, which strongly suggests that the standard LOD prediction using indentation alone is not an accurate method for evaluating the global importance of lines of Java code.

Aside from indentation, the most obvious information about the source code is the type of programming construct. If construct type can predict the LOD of a given line, it would be expected that certain types would attract LOD ratings that differed significantly from other types. This was observed to be true, and highly significant differences were revealed between the LODs of different constructs. The source of this variation was from 'end' types and from simple 'statement' types, which differed significantly both from each other, and from all the other types, with a much lower assessed LOD.

It was hypothesized that these two construct types may have been affecting the original algorithm which used indentation as a measure of LOD. The analysis was repeated, this time excluding the data from the 'end' and 'statement' types. This time, the results were shown to be highly significant, indicating that there were considerable differences between the LOD values assigned to the different levels of indentation. Furthermore, more detailed analysis to determine the nature of this variation revealed a highly significant linear correlation between indentation and genuine LOD.

This evidence suggests that the LOD calculation used to generate the fisheye view for Java source code might be improved with only minor alterations to the original algorithm, by excluding 'end' and 'statement' construct types from the calculation and assigning them lower values for LOD, independent of their level of indentation.

There are however, two issues that need to be considered:

1. This experiment only dealt with six differing levels of indentation (five when excluding simple statements). The mean LOD value for a 'statement' type was 4.25, while, at the most indented level, the mean LOD value for 'other' types was 4.4. Increasing levels of indentation may mean that a degree of overlap arises between these types. Further experimentation with more deeply nested code samples is required to determine whether this 'overlapping' does in fact occur and to establish how to address this issue.

2. The second issue is that 'statement' and 'end' type constructs are by far the most common type of construct encountered in code (in this study they comprised 49% and 22% of all the lines, respectively). It is likely that the LOD of lines of code within these groups may vary, and additional measures may be needed to differentiate between the LOD values of 'simple statement' and 'end' types as a whole.

Analysis showed that neither the LOD of 'end' types nor of 'simple statement' types could be predicted by indentation. Whilst it may be sufficient to assign identical LOD values to all 'end' type code lines, for 'simple statement' types there was a much higher variation between the LODs of individual lines.

The final analysis indicated that the importance of statement types might be associated with their sub-class. It revealed significantly different LOD values between all four statement sub-types ($p<0.05$). This would suggest that statement sub-class may be used to determine LOD, with 'Method Calls' being assigned the lowest LOD, followed by 'Other' types, 'return' statements and lastly by Declaration/ Assignment types.

It is worth noting that, although significant differences were observed between these four statement types, this does not mean that this is necessarily the best method to use to differentiate between different types of statement, or indeed that these are the only four possible sub-types.. Further investigation is needed to look specifically at how best to calculate the LOD of 'simple statement' types to most accurately reflect their genuine LOD.

This division of 'simple statement' types may also lead to a degree of overlap between the LODs of Declaration/Assignment types and the most indented constructs. This would also need to be investigated more fully, particularly with regards to increasingly deeply nested structures, as mentioned previously.

6 Conclusions and Further Work

The results of this study indicate that there may be some truth in the notion that LOD can be determined by the level of indentation. However, they also suggest that it may not be most accurate or effective method of LOD determination for all Java construct types.

The findings of the study strongly suggest that, for 'end' and 'simple statement' types at least, indentation level has no correlation with global importance. Simply by excluding these construct types from the standard LOD calculation we can markedly improve the accuracy of the LOD determination for the remaining types. Further research is needed to determine how much this simple alteration to the LOD determination algorithm can improve the generated fisheye view.

The purpose of this study was to determine whether a more accurate, easily available LOD calculation could be found *for Java source code*. The results we obtained may not apply to other programming languages (particularly non object-oriented languages). For other domains using indent structuring (such as other programming languages, biological taxonomies or decision trees) the general LOD algorithm originally developed by Furnas [5, 6] may indeed prove to be the most accurate method.

Nevertheless, for the purposes of this particular project, it would appear that this updated method of LOD calculation will match more closely the user's opinion of the importance of a given line of code. This may be sufficient to improve the generated fisheye view to such a level whereby the lines displayed by default more closely match the user's expectations with less need for a manual override control.

Acknowledgements. This work was undertaken as part of a project being supported by a studentship awarded by the University of Aberdeen.

References

1. Cockburn, A.: Supporting Tailorable Program Visualisation Through Literate Programming and Fisheye Views. Information and Software Technology 43, 745–758 (2001)
2. Cockburn, A., Smith, M.: Hidden Messages: Evaluating the Efficiency of Code Elision in Program Navigation. Interacting with Computers. The. Interdisciplinary Journal of Human-Computer Interaction 15(3), 387–407 (2003)
3. Finlayson, J.L., Mellish, C.: The AudioView - Providing a Glance at Java Source Code. In: Proceedings of ICAD'05 - Eleventh meeting of the International Conference on Auditory Display (2005)
4. Foltz, P.W., Landauer, T.K., Parker, J.: Design principles for organization and navigation in interactive electronic technical manuals: A review of relevant science and technology (2001) http://www.k-a-t.com/ONR/ONR-HT-IETreview.PDF
5. Furnas, G.W.: The FISHEYE view: a new look at structured files. Bell Laboratories Technical Memorandum #81-11221-9 (1981)
6. Furnas, G.W.: Generalized Fisheye Views, Human Factors in Computing System. In: Proceedings of CHI'86, pp. 16–23 (1986)
7. Schaffer, D., Zuo, Z., Greenberg, S., Bartram, L., Dill, J., Dubs, S., Roseman, M.: Navigating hierarchically clustered networks through fisheye and full-zoom methods. ACM Transactions on Computer-Human Interaction 3(2), 162–188 (1996)
8. Virpioja, S.: Fisheye minor mode for emacs (2005) http://www.niksula.hut.fi/s̃virpioj/fisheye

A Bayesian Network Approach to Semantic Labelling of Text Formatting in XML Corpora of Documents

Florendia Fourli-Kartsouni, Kostas Slavakis, Georgios Kouroupetroglou, and Sergios Theodoridis

Department of Informatics and Telecommunications
University of Athens, GR-15784, Athens, Greece
{fkartsou,slavakis,koupe,stheodor}@di.uoa.gr

Abstract. The wide-spread applications of document digitization have lead to the use of structured digital representation methods such as the XML language. Extraction methodologies for the formatting metadata can be used on such structured documents for enhancing their accessibility, including augmented audio representation of documents. To the best of our knowledge, an effort has yet to be made to produce an automatic extraction system of semantic information of the document formatting, solely from document layout, without the use of natural language processing. In this study a corpus of XML representations of several issues of a Greek newspaper is used in order to create and evaluate a semantic classifier of text formatting, based on Bayesian Networks.

Keywords: document accessibility, document analysis, semantic labeling.

1 Introduction

As document digitization spreads out rapidly many traditional sources of information such as books, newspapers and journals tend to appear more and more in digital form. Digital information has many advantages as it requires less physical storage space, it is searchable, easier to transmit and receive, more manageable and significantly more accessible. In particular, document digitization has an important impact on information accessibility for the visually impaired and print disabled readers. Not only they can have access to a larger amount of information, but also this information can be structured, and hence, be used more effectively [22]. Moreover, audio access to documents is becoming popular for use in wider types of settings, for example while driving, using a mobile phone or as part of a learning course [20], [21].

Digital representation can incorporate several levels of the structure of documents such as the physical, and the logical layout structure. The physical structure level corresponds to the physical form of a document, i.e. zoned into pages, regions of pages, blocks of text and so forth. The logical structure level is associated with the documents layout interpretation. For example, a block of text in the physical structure level might be interpreted as a title in the logical level.

Digital representation methods for documents vary. The representation of the physical characteristics of the document can be carried out using methods of a

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 299–308, 2007.
© Springer-Verlag Berlin Heidelberg 2007

complicated structure, such as the XML language. A digital document based on an XML format allows representation of the documents' physical and logical structure, while preserving presentation information and can even expose semantics. Using such a representation we pass from the level of data to the level of meta-data.

The primary focus of this study is the extraction of semantic information conveyed in the documents' physical characteristics described in XML documents. For our study, we used a corpus of XML representations of several issues of a Greek newspaper. These XML documents use meta-data describing the documents physical layout, including textual features (for example, type weight and style), a number of geometrical features and most of the documents' elementary logical structure (for example, article, title, subtitle, and paragraph).

Often, textual features present purely logical structure information. For example, a piece of text can be part of a list or footnote. However, many times the use of these features has the purpose of communicating semantic information to the reader. Examples of such purposes are: giving emphasis to an important piece of text; focusing the readers' attention to the central point of an article; highlighting a name i.e. the title of a movie and so on. This classification can help the creation of sophisticated document manipulation tools including an enhanced audio representation of the document.

To the best of our knowledge, this is the first time of an attempt to produce an automatic extraction system of semantic information based only on the document layout, without the use of natural language processing. However, there have been a number of studies on the automatic identification of logical structure of documents e.g. [1], [2], [4], [8]. Most traditional approaches in this field have employed deterministic methods (decision trees, formal grammars) [12], [15], [4], which however may suffer from poor performance in the presence of noise and uncertainty. In addition, such approaches create models which are not flexible to domain changes and cannot easily evolve in the presence of new evidence.

In order to overcome such limitations we employ a probabilistic approach based on Bayesian networks trained on a series of labelled documents. Bayesian networks offer a significant tolerance to noise and uncertainty, they can capture the underlying class structure residing in the data and they can be trained on examples, thus adapting to existing and future evidence. Models of this type have been previously employed for document and natural language processing tasks such as logical labelling [14], [13], [19], Part-Of-Speech (POS) tagging [10], character segmentation in hand written text [9] and many more and have showed to outperform deterministic equivalents.

In the following sections we present an attempt of identifying the semantics conveyed in the use of textual features (such as *bold* and *italics*), a methodology for automatic semantic labelling of text, based on these features and some preliminary results.

2 Semantic Interpretation of Text Formatting

Within the field of typography there exists a plethora of guidelines for text formatting (e.g. [3], [11], [17] and [18]). Guidelines present great variation depending on the type of publication (for example book or journal), the presentation medium (for

example printed or online) and each particular publisher. This imposes a great complication to the attempt of collecting a concrete set of semantic interpretation of text formatting. Moreover it is not always safe to assume that the intention of the author in using a particular set of textual features is always correctly understood by the reader.

In an effort to simplify the task at hand, we focused our study only on two basic textual features, namely the weight (normal or bold) and the style of type (normal or italics). In addition, we drew our set of semantic interpretations on the preliminary conclusions of an ongoing survey, which exposed readers to various text formatting examples selected from our corpus. This ensured that we don't limit our study on a compilation of rules from various guidelines, thus ignoring the reader's interpretations and limiting our model to a deterministic rule based system.

2.1 Interpretation by the Readers

We are currently conducting a survey for identifying reader interpretation of text formatting. The methodology is based on presenting the readers with examples of formatted text and then asking them to describe the purpose that they believe the specific text formatting serves.

So far, we have collected and processed the responses of a total of thirty-five readers of varying age groups, education levels and other demographic characteristics. The subjects were presented with a number of short parts of text taken from newspaper articles. The articles used were all in Greek, the native language of all the readers. The cases spanned through various examples of the use of “bold” and “italics” type in single words, phrases or sentences. The readers were asked to note down, next to each part of text, their semantic interpretation using the following question as a guide: “What do you think was the intention of the writer for applying the particular formatting characteristic to this text?”

2.2 Set of Semantic Labels

Based on the statistical results from our survey, we managed to identify eight different “labels” that the readers seem to use most frequently in order to semantically characterize text with “bold” and “italics” formatting:

Emphasis: A word or phrase that is considered significant and needs to be stressed out.
Important / Salient: A word or phrase, which is near or is part of a piece of information that is considered important and should be noticeable. A word or phrase that “catches the eye” of the reader.
Basic Block: A block of text, which introduces or summarizes the main content of the article.
Quotation: A piece of text corresponding to a fragment of written or oral expression of a person other than the writer of the article.
Note: A piece of text serving at providing additional information or explanation related to part or to the whole of the article.

Title: A piece of text corresponding to the name of a movie, play, book and so on or the title of a newspaper, television channel or journal.
List / Numeration Category: A word of phrase that is part of a list or a numeration and serves as a new "instance" indicator.
Interview / Dialogue: A piece of text that is part of an interview (the question or the answer) or that corresponds to a dialogue between two persons.

A few *"logical"* labels, such as "subtitle" or "footnote" were also mentioned by some readers but these were not considered to be *"semantic"* labels and were therefore not of value to the purposes of this study.

3 Classification Features and Methodology

The XML files in our corpus describe a great deal of the documents' structure, including the logical structure by using meta-data tags such as *<article>, <title>, <paragraph>* and so on. Text formatting is also denoted with the use of such tags as *<b>* for "bold", *<i>* for italics, common in markup languages. The XML files were parsed in order to maintain only the text which appeared within "bold" and "italics" tags. The piece of text within every occurrence of these tags was considered an *entity* to be classified into one of the semantic categories as described in section 2.2.

Further textual parsing of each entity allowed the extraction of a number of additional features and the representation of the entity as a vector used for the training and evaluation of the classifier.

The characteristics employed for the representation of the physical structure of text entities include textual, geometrical and neighborhood features and finally the presence or absence of some special characters. The chosen representation captures the combination of the set of characteristics. It can describe, for example, *a long block of text which is bold, it is at the beginning of a paragraph and is preceded by a text block ending with a question mark.*

In total, each vector representing an entity consists of the values of the following seventeen features:

1. *Bold:* This feature corresponds to an entity being inside a "bold" tag (a value of 1) or not (a value of 2).
2. *Italics:* This feature corresponds to an entity being inside an "italics" tag (a value of 1) or not (a value of 2).
3. *Top:* This feature corresponds to the entity being found at the top of the article (a value of 1) or not (a value of 2).
4. *Long:* This feature corresponds to an entity consisting of more than 20 words (a value of 1) or not (a value of 2).
5. *Short:* This feature corresponds to an entity consisting of less than 3 words (a value of 1) or not (a value of 2).
6. *OneWord:* This feature corresponds to an entity consisting of only one word (a value of 1) or not (a value of 2).
7. *Quotes:* This feature corresponds to an entity including an opening or closing ('«' or '»') quote character (a value of 1) or not (a value of 2).

8. *Numbers:* This feature corresponds to an entity including an arithmetic character (a value of 1) or not (a value of 2).
9. *QuestionMark:* This feature corresponds to an entity including a "question mark" character (a value of 1) or not (a value of 2).
10. *Dot:* This feature corresponds to the entity including a dot character (a value of 1) or not (a value of 2).
11. *FirstCapital:* This feature corresponds to an entity starting with a capital letter character (a value of 1) or not (a value of 2).
12. *PreviousQuote:* The value of feature Quote of the entity preceding the one under examination.
13. *FollowingQuote:* The value of feature Quote of the entity following the one under examination.
14. *PreviousTop:* The value of feature Top of the entity preceding the one under examination.
15. *FollowingQuote:* The value of feature Top of the entity following the one under examination.
16. *PreviousQuestionMark:* The value of feature QuestionMark of the entity preceding the one under examination.
17. *FollowingQuestionMark:* The value of feature QuestionMark of the entity following the one under examination.

3.1 Bayesian Networks

Bayesian networks are directed acyclic graphs, which encode the *Markov Assumption* and allow for an efficient representation of the joint probability of a set of random variables.

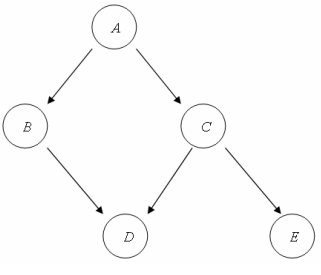


Fig. 1. A simple Bayesian network structure

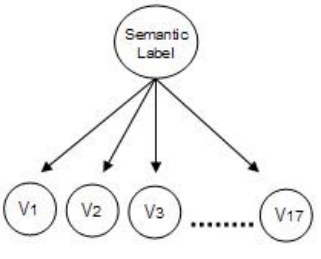


Fig. 2. The Naive Bayes Classifier

Each vertex in a Bayesian network represents a random variable, and edges represent dependencies between the variables. For example, in Fig.1 variable B is dependant on variable A. Moreover the fact that the arrow points from A to B makes A a *parent* of B and B a *descendant* of A.

Now consider a random variable $X_i \in \{X_1, X_2, ...X_n\}$, $1 \le i \le n$, which in our case takes on values x_i from a finite and discrete set of real numbers. It is

common to denote $P(X_i = x_i)$ as the probability of X_i assuming one of its possible values x_i.

The *joint probability* of a set of random variables $\{X_1, X_2, ...X_n\}$ can be defined for formally as a function $P(X_1, \ldots, X_n)$, which should always satisfy:

$$0 \le P(X_1, \ldots, X_n) \le 1, \text{ for every combination of } x_i, \tag{1}$$

and

$$\sum_{x_1, \ldots, x_n} P(X_1, \ldots, X_n) = 1$$

The *conditional probability* of X_i, given a second variable X_j, is the probability of X_i conditioned on the fact that X_j assumes a value x_j. It is denoted as $P(X_i = x_i \mid X_j = x_j)$ and is defined as:

$$P(X_i = x_i \mid X_j = x_j) := P(x_i \mid x_j) := \frac{P(x_i, x_j)}{P(x_j)}. \tag{2}$$

The structure of this network implies several conditional independence statements. We say that two random variables X_i and X_j are conditionally independent given a third random variable X_z when:

$$P(X_i \mid X_j, X_z) = P(X_i \mid X_z). \tag{3}$$

For example in Fig. 1, variable C and B are conditionally independent given A, which, is denoted as: $I_P(\{C\}, \{B\} \mid \{A\})$.

The *Markov Assumption* states that each variable X_i is independent of its non-descendants ND_{X_i} in the graph given the state of its parents PA_{X_i}. We formally write this as:

$$I_P(\{X_i\}, ND_{X_i} \mid PA_{X_i}), \forall i. \tag{4}$$

This assumption allows us to use (3) in order to derive that for Bayesian networks:

$$P(X_i \mid ND_{X_i}, PA_{X_i}) = P(X_i \mid PA_{X_i}) \tag{4.1}$$

Using (2) for calculating the joint probability distribution and by simply applying the common *chain rule* it is easy to derive:

$$P(X_1, \ldots, X_n) = P(X_n \mid X_{n-1}, \ldots, X_1) \cdots P(X_2 \mid X_1) P(X_1) \tag{5}$$

Based on the Markov Assumption we can then use (4.1) to simplify the above calculation even more:

$$P(X_1,\ldots,X_n) = P(X_n \mid PA_{x_n})P(X_{n-1} \mid PA_{x_{n-1}})\cdots P(X_1 \mid PA_{x_1}) \quad (6)$$

Based on the above simplification, Bayesian networks reduce the number of parameters needed to characterize a joint probability. Moreover, they enable the encoding of "a-priori" knowledge and causal relationships in the model, and facilitate efficient computation of posterior probabilities given evidence ([16]).

3.2 The Final Model

The final model to be used as a semantic classifier for text formatting was based on a Bayesian network which is commonly known as the *Naive Bayes Classifier* [6], [7] depicted in **Fig. 2**. This network is based on the "naive" assumption that every feature (every leaf in the network) is independent from the rest of the features, given the state of the class variable.

The network consists of 17 nodes, V_1 to V_{17} representing the seventeen features used for describing the text formatting of entities presented in chapter 3 and the "*Semantic Label*" root node representing the class variable which can take values from the 8 labels identified in chapter 2.2.

Consider a previously unseen instance from a labelled training sample $\mathrm{I} = \{ v_1, v_2, \ldots, v_{17} \}$ where v_i is the value of the i^{th} feature V_i. The model predicts the class label for I by assigning it to c that belongs to the set that maximises $P(c_i \mid v_1, v_2, \ldots, v_{17})$, that is:

$$\arg\max_{c_i}(P(c_i \mid v_1, v_2, \ldots, v_{17})) = \arg\max_{c_i}(P(c_i)P(v_1, v_2, \ldots, v_{17} \mid c_i)) \quad (7)$$

Where $c_i \in \{c_1, c_2, .., c_8\}$ is the i^{th} value of the class variable. Taking into consideration the independence assumptions in our model, (7) is simplified to:

$$\arg\max_{c_i}(P(c_i)\prod_{j=1}^{17} P(v_j \mid c_i)) \quad (7.1)$$

The class probability $P(c_i)$ is easily calculated by counting the occurrences of each label in the training set and dividing by the total number of training instances. The estimation of the conditional probabilities $P(v_j \mid c_i)$, which maximise the likelihood of the training data, can also be achieved by "counting". This time the occurrences of v_i in data examples labelled as c_i are counted and divided by the total number of occurrences of v_i in the training data.

This simple and straightforward model is widely used for classification problems and has proved surprisingly successful even though it does on some cases suffer due to sparse data or variables dependencies [5].

4 Results

The Naive Bayes semantic classifier was constructed and evaluated using a corpus of 2.000 articles of a Greek newspaper in XML form. A total of 2.927 entities, of which 1866 were occurrences of "bold" and 1061 of "italics" were manually labelled based on the finding of our survey. In order to better estimate the performance of our classifier we adopted a 10-fold cross validation method.

The corpus was partitioned in ten equal parts. The 90% of those parts was used for training the classifier while the remaining 10% was used for evaluating its performance. This process was repeated ten times including different entity examples for the two parts at each cycle. The overall accuracy of the classifier was measured by averaging the test results of all the cycles.

The total accuracy of the classifier at each cycle was measured by counting the correctly classified test examples and dividing by the total number of examples in the test set. In addition classification precision and recall were measured for each of the 8 labels. Table 1 summarises the overall results.

Table 1. Naïve Bayes classifier results

		Total Accuracy: 89%
Precision	Recall	Class
0.98	0.979	Basic Block
0.965	0.961	Emphasis
0.947	0.831	Importance / Salience
0.971	0.95	Interview / Dialogue
0.864	0.84	List
0.727	0.851	Title
0.6	0.578	Quotation
0.577	0.668	Note

As we can see from the results the classifier has an overall satisfactory performance. However the performance is significantly lower for the last three labels namely "Title", "Quotation" and "Note". Looking at the graph at Fig. 3 we can see the distribution of the data set examples to the various labels.

It is obvious from this chart that the labels that the classifier has the lowest performance on are also the ones with the fewer samples in the dataset. This can partially explain the variety in classification precision as it is a known weakness of Naïve Bayes classifiers. However after some further investigation we discovered that the labels on which the model does worst on are those that correspond to mostly "italics" text. Looking at the confusion matrix of the classification (Table 2) we can see numerous occurrences of misclassification of the three labels but not of label "List" that also has few samples in the dataset.

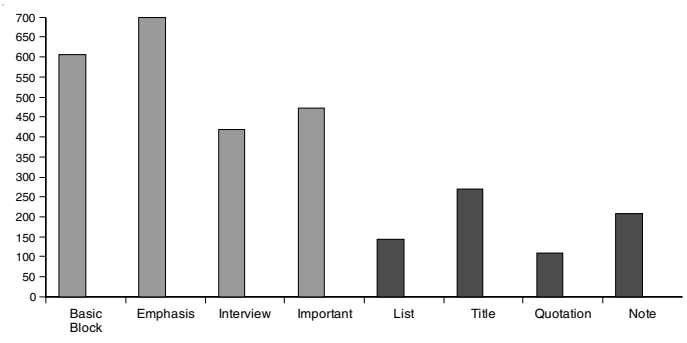


Fig. 3. Sample distribution to the semantic labels

Table 2. Classification Confusion Matrix

A	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*classified as*
93	5	0	4	3	0	1	0	A:BasicBlock
2	671	6	8	2	3	6	0	B: Emphasis
2	5	399	1	2	9	2	0	C: Interview
2	10	1	121	2	7	0	1	D: List
0	1	1	3	139	11	41	12	E: Note
6	1	2	0	22	63	10	5	F: Quotation
0	2	0	3	19	12	229	4	G:Title
0	0	2	0	52	0	26	392	H: Importance

This observation lead us to conclude that the labels that correspond to “italics” text are not clearly defined and separated using the set of features selected for this study. Overall the performance of the Naïve Bayes classifier was quite satisfactory but more sophisticated models and additional text formatting features are sure to be examined in future research.

Acknowledgements. The work described in this paper has been funded by the European Social Fund and National Resources under the RHETOR project of the Information Society Operational Programme: Image, Sound and Language Processing, Greek General Secretariat of Research and Technology.

References

1. Conway, A.: Page grammars and page parsing: a syntactic approach to document layout recognition. In: Proc. Int. Conf. on Document Analysis and Recognition, pp. 761–764 (1993)
2. Yamashita, A., Amano, T., Takahashi, I., Toyokawa, K.: A model based layout understanding method for the document recognition system. In: Proc. Int. Conf. on Document Analysis and Recognition, Saint Malo, France, pp. 130–138 (September 1991)
3. Chicago Manual of Style. 15th edn. Chicago: University of Chicago Press (2003) http://www.chicagomanualofstyle.org/

4. Derrien-Peden, D.: Frame-based system for macro-typographical structure analysis in scientific papers. In: Proc. Int. Conf. on Document Analysis and Recognition, Saint-Malo, France, pp. 311–319 (1991)
5. Rish, I., Hellerstein, J., Jayram, T.: An analysis of data characteristics that affect naive Bayes performance. Technical Report RC21993, IBM Watson Research Center (2001)
6. Langley, P., Iba, W., Thompson, K.: An analysis of Bayesian classifiers. In: Proc. 10th Nat. Conf. Artificial Intelligence, pp. 399–406. AAAI Press and MIT Press, Stanford, California, USA, Cambridge, MA (1992)
7. Langley, P., Sage, S.: Induction of selective Bayesian classifiers. In: Proc. 10th Conf. Uncertainty in Artificial Intelligence, pp. 223–228. Morgan Kaufmann, San Francisco (1994)
8. Krishnamoorthy, M., Nagy, G., Seth, S., Viswanathan, M.: Syntactic segmentation and labeling of digitized pages from technical journals. IEEE Transactions on Pattern Analysis and Machine Intelligence 15, 737–747 (1993)
9. Maragoudakis, M., Kermanidis, K., Fakotakis, N., Kokkinakis, G.: Combining bayesian and support vector machines learning to automatically complete syntactical information for HPSG-like formalisms. In: Proceedings of International Conference on Language Resources and Evaluation, Las Palmas, Spain, pp. 93–100 (2002)
10. Maragoudakis, M., Ganchev, T., Fakotakis, N.: Bayesian reinforcement for a probabilistic neural net part-of-speech tagger. In: Proc. Int. Conf. on Text Speech and Dialogue, Brno, Chech Republic, pp. 137–145 (2004)
11. Bringhurst, R.: The Elements of Typographic Style, 2nd edn. pp. 93–119. Hartley & Marks Publishers, Vancouver Canada (2002)
12. Mao, S., Rosenfeld, A., Kanungo, T.: Document structure analysis algorithms: a literature survey. In: Proceedings of SPIE 5010, pp. 197–207 (2003)
13. Souafi-Bensafi, S., Parizeau, M., Lebourgeois, F., Emptoz, H.: Bayesian networks classifiers applied to documents. In: Proc. IEEE ICPR vol. 1, pp. 483–486 (2002)
14. Souafi-Bensafi, S., Parizeau, M., Lebourgeois, F., Emptoz, H.: Logical labeling using bayesian networks. In: Proceedings of IEEE ICDAR, pp. 832–836 (2001)
15. Tsujimoto, S., Asada, H.: Understanding multi-articled document. In: Proc. Int. Conf. on Pattern Recognition, Atlantic City, NJ, pp. 551–556 (1990)
16. Theodoridis, S., Koutroumbas, K.: Pattern Recognition, 3rd edn. pp. 13–26. Academic Press, San Diego (2006)
17. The American Psychological Association. Publication Manual, Washington DC, pp. 94–103, 111–130 (2001)
18. The Economist Style Guide http://www.economist.com/research/StyleGuide/
19. Tateisi, Y., Itoh, N.: Using stochastic syntactic analysis for extracting a logical structure from a document image. In: Proc. Int. Conf. on Pattern Recognition, Israel, pp. 391–394 (1994)
20. Xydas, G., Argyropoulos, V., Karakosta, T., Kouroupetroglou, G.: An Experimental Approach in Recognizing Synthesized Auditory Components in a Non-Visual Interaction with Documents. In: Proc. 11th Int. Conf. Human-Computer Interaction, Las Vegas, pp. 411–420 (2005)
21. Xydas, G., Kouroupetroglou, G.: Augmented Auditory Representation of e-Texts for Text-to-Speech Systems. In: Matoušek, V., Mautner, P., Mouček, R., Tauser, K. (eds.) TSD 2001. LNCS (LNAI), vol. 2166, pp. 134–141. Springer, Heidelberg (2001)
22. Web Accessibility Initiative (WAI) http://www.w3.org/WAI/

Winsight: Towards Completely Automatic Backtranslation of Nemeth Code

D. Gopal[1], Q. Wang[1,2], G. Gupta[2], S. Chitnis[1], H. Guo[3], and A. Karshmer[4]

[1] Logical Software Solutions, Plano, TX
[2] University of Texas at Dallas, Richardson, TX
[3] University of Nebraska at Omaha, Omaha, NE
[4] University of San Francisco, San Francisco, CA

Abstract. We present the Winsight system, a Windows-based software system for completely automatic translation of Nemeth Braille code to LaTeX. The Winsight system takes hard copy Braille input containing Mathematics (written in Nemeth Braille code) and text (written in contracted Braille) via a scanner, performs image recognition and analysis of the scanned file to generate the ASCII Braille file, automatically separates Nemeth Braille coded expressions and contracted Braille text, backtranslates them to LaTeX math expressions and LaTeX text respectively, and produces a print output file in pdf format containing the result of backtranslation. The Winsight system comes with tools that allow users to manually intervene during each step, if they desire, to fix any errors reported by the system or seen by the user. In this paper we give an overview of the Winsight system.

1 Introduction

Visually Impaired individuals "read" and write using Braille. Various encodings have been used to make Braille writings more succinct and efficient to read (for example, contracted Braille codes). Braille-based notations have also been developed to encode Mathematical expressions, for example, the Nemeth Braille code for mathematics [4]. Visually impaired students in the United States (as well as in other countries such as New Zealand, Canada, Greece, India, etc.) typically learn the Nemeth Math notation to "read" and write Mathematics. However, when they take a Mathematics course, the following problem arises. The instructor (who is typically sighted) prepares all the course material (such as home assignments, exams) using LaTeX, a popular word processing package used for typesetting Mathematics [7]. The visually impaired student cannot access this homework; the homework has to be read out to him by someone. Next, the visually impaired student might write the answers to problems in the homework using Nemeth Math notation. The solution is unreadable by the instructor since he/she neither knows the Nemeth notation nor Braille. This communication gap between visually impaired students and their sighted instructors is a major problem that hinders Mathematics and Science education for visually impaired

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 309–318, 2007.
© Springer-Verlag Berlin Heidelberg 2007

individuals. In this paper we describe our automatic solution for bridging this communication divide.

To illustrate the problem, consider how visually impaired students communicate with their sighted instructors. For example, in response to a question, a visually impaired student may write the following proof using contracted Braille and Nemeth Math Braille code (we use the ASCII equivalent of the Braille code called ASCII Braille; ASCII Braille code is normally used to represent Braille on a computer):

```
,pro( ,let ;a & ;b be two po9ts 9 ! 9t]val )
a "k b !n "! is "s x 9 (a, b) )
    f'(x) .k ?f(b)-f(a)/b-a#,
b f'(x) .1 0 = all x 9 (a, b) ,so
      ?f(b)-f(a)/b-a# .1 0,
s9ce b-a .1 0 x foll[s t f(b) .1 f(a)
```

As the reader can notice, this is hard to read without knowledge of the two Braille codes (contracted Braille and Nemeth code) involved. The above proof stands for:

Proof Let a and b be two points in the interval with $a < b$ then there is some x in (a, b) with

$$f'(x) = \frac{f(b) - f(a)}{b - a}$$

But $f'(x) > 0$ for all x in (a, b) So

$$\frac{f(b) - f(a)}{b - a} > 0$$

Since $b - a > 0$ it follows that $f(b) > f(a)$

One way to solve this communication problem is to build automatic translators between LaTeX and Nemeth Math notation so that the homework assignment can be automatically translated from LaTeX to Nemeth code for the student to "read," while the homework answers prepared by the student in Nemeth code can be automatically translated from Nemeth Math to LaTeX for the sighted instructor to read (the process of translating Nemeth code to LaTeX is called *backtranslation*). Building the translator and backtranslator, however, is a difficult and time-consuming task.

Building a translator that automatically translates LaTeX to Nemeth Math is relatively easy, as LaTeX is a recently designed notation, and has a *context free grammar* (which means that LaTeX documents can be easily parsed automatically by a computer). Indeed, a translator from LaTeX to Nemeth Math is commercially available from MacKichan Software as part of their Scientific Notebook System [8] (currently marketed by Duxbury Systems). In contrast, the Nemeth Braille Math notation is quite complex, since it is designed to be human readable. As a result, the backtranslation problem is considerably more difficult and in the past has been considered as unsolvable [6]. The problem of backtranslation becomes

harder if Nemeth code is embedded within contracted Braille text. In fact, most mathematical documents will contain both Math expressions and English text.

Assuming that a backtranslator system is available, an additional problem arises. The backtranslator will take the Math document written by blind individual as input, which has to be provided in an electronic form, such as ASCII Braille. Typically most blind students write mathematical documents using a mechanical Brailler (e.g., Perkins Brailler). Electronic Braille devices (such as *Notetakers*, with refreshable Braille display) which can produce an electronic Braille document are not used even if a blind person has access to them. This is because such Braille displays can only display one line at a time, and while solving mathematical problems, frequently the writer needs to refer back to several of the previous lines, especially, if spatial structures (e.g., matrices, determinants) are used. The mechanical Braillers only produce a hard-copy document, there is no electronic file produced. The problem then one faces is the conversion of hard-copy Braille documents into an electronic format that can be fed to the backtranslator.

In this paper we report on our work on the Winsight system, a system that *completely automates* the backtranslation process. The system consists of 3 modules: the OBR module, the *separation module* and the *backtranslation module*, all developed by us. Hard copy Braille documents produced by mechanical Braillers and containing both Math (written in Nemeth Braille Code) and English text (written in level 2 Braille) are scanned on a standard scanner, and the image file obtained automatically processed by the Winsight OBR module to recognize the Braille characters. The Braille characters recognized are stored in a file in the ASCII Braille format. This file is further processed by the Winsight separation module to automatically detect parts that are written in Nemeth code and those that are written in contracted Braille. This prior separation of Nemeth and contracted Braille components is important, because, in general, it is difficult to automatically distinguish between the two codes (for example, the symbol '+' means different things in the two codes). The separation module will insert the string `;;` at the beginning and the string `;'` at the end of each Nemeth math expression it recognizes. Text outside of these begin-end symbols is then assumed to be in contracted Braille. This annotated file is then fed to Winsight's backtranslation module, which produces a LaTeX file that can then be rendered in the pdf format to be read by a sighted person (e.g., a course instructor).

An additional feature of the Winsight system is that the user can intervene after any of the intermediate steps to fix problems. Thus, the user can edit the image file after it has been scanned. He/she can edit the *blob* file obtained after image processing, as well as the ASCII Braille file obtained after the Braille character inferencing has been done. Finally, the user can also edit the result of automatic separation of Math and contracted Braille text. This user intervention is needed because typically hand typed Braille documents have many errors of spacing (e.g., paper may have moved during typing), the Braille character may have been intentionally erased by the user either by backspacing or by typing

the '=' Braille character over mistyped text, etc. Additionally, the Nemeth Math code typed may have syntactic errors.

2 Nemeth Math Braille Code

Nemeth Braille Math notation (hereafter referred to as Nemeth code) is a Braille-based notation to enable visually impaired scholars and students to read and write Mathematics [4]. It was designed by Dr. Abraham Nemeth in the late 1940s to allow visually impaired individuals to convey mathematical and technical information. While Nemeth code is well suited for humans, it is not easy to process with a computer. To make Nemeth code computer processable we should be able to *parse* Nemeth code text for grammatical correctness (i.e., check that it conforms to the rules of syntax laid out in the Nemeth code specification [4]), and once parsed, transform it into a form so that it can be processed further. To the best of our knowledge there are no parsers available that will parse Nemeth code to check its grammatical correctness. Two principal reason for this lack of availability of parsers are: (i) context sensitivity, and (ii) spatially arranged mathematical structures.

Context Sensitivity: The Nemeth code notation falls in the class of languages known as *context sensitive languages* [1]. Context sensitivity means that the grammatical correctness of a subexpression depends on the context in which this subexpression occurs within a larger expression. This is similar to the case of English, where, for example, to check the grammatical correctness of the sentence one has to make sure that a singular subject noun must have a singular verb. So `The man eats an apple` is grammatically correct, while `The man eat an apple` is not. Another negative point about the Nemeth code is that it is not an *LALR(1)* [1] language. This means that if we were to write grammar rules for Nemeth code, then given a Nemeth code expression that we are trying to parse, we cannot tell by looking ahead one symbol in the expression which grammar rule should be applied to parse the expression. Thus, a grammar written for a LALR parsing tool such as `YACC` will generate too many *conflicts* [1] at the time of parsing. As a result, traditional parsing and compiling technology [1], that is normally used to process computer languages, cannot be used.

It should be noted that properties of the Nemeth code—especially context sensitivity—are precisely the reason why Nemeth code is so suitable for a visually impaired person. Since a blind person reads by touching the Braille code, it is important that he/she be aware of the context of the current symbol or expression (s)he is currently looking at, at any given time. Thus, building-in the context into the language is quite helpful for blind "readers," but makes its computer processing quite difficult. For example, the expression $x^{z+2^c+3} + y$ is typeset as `$x^{{z+2}^{c}+3}+y$` in LaTeX, since to a sighted reader (or a computer) the braces indicate the scope of exponents. However, this does not work well in a Braille setting, as the reader will quickly forget the exponent level as he/she moves left to right in the formula; considerable number of backtrackings of the finger will be needed to understand it. In Nemeth code, to make the

context explicit, this formula is coded as: `x^z+2^^c^+3"+y`. This makes parsing hard, as a context free grammar that can count the number of ^ and check that their number increases or decreases by 1 as we go up and down the exponent level cannot be written [1].

Also, at the time the Nemeth code was designed (late 40s and early 50s), computers were just being invented. Thus, computer readability or computer processability was not one of the design criteria for Nemeth code. Additionally, very little was known about structure of formal languages then. It is very likely, the Nemeth code was designed with the intention that it will be translated into print by humans, who can easily process context sensitivity. Recently, techniques have been developed by Gupta, Guo, and Karshmer—incorporated in the Winsight system—that allow one to build a parser very rapidly and transform Nemeth code into LaTeX [3,2].

Spatial Arrangements: Another feature of the Nemeth code that makes it hard to parse are its spatial arrangements of Mathematics [2]. There are mainly two kinds of spatial arrangements in Nemeth code. The first is the spatial arrangements for arithmetic operations such as addition, subtraction, multiplication, and division (including synthetic division), and square-root. The second is the spatial arrangements for matrices, determinants, hypercomplex fractions, and unified expressions. Parsing spatial arrangements is a difficult task.

The backtranslator that is part of the Winsight system parses spatial arrangements by analyzing them several lines at a time to look for patterns within horizontal lines as well as along vertically aligned characters. It also identifies boundaries of non-spatially arranged mathematical expressions that are part of the spatial structure (e.g., a polynomial element inside a matrix). Winsight then translates the non-spatially arranged structures into LaTeX. These LaTeX translated expressions are then arranged by Winsight into the appropriate spatial structure using LaTeX's constructs and then output.

The representation of spatial arithmetic in Braille is guided by numerous rules [4]. These rules have to be used as guidelines in the spatial reasoning module of the Winsight system to accomplish the translation of spatially arranged arithmetic and algebra.

In simple addition, the units digit, tens digit, etc., should always be aligned. Thus, in Nemeth code the addition of two numbers 25 and 125 can be represented only in the following way:

```
  1
77777
  25
+125
33333
 150
```

Notice that 5 (the units digit) in 25 is vertically aligned with 5 in 125 and so on. The operator '+' or '-' should always be present at least one column to the left of the longest element in the operation. The "333333" notation is used as

the separator line below which the answer is written. The length of this operator column is at least one greater than the largest row. The carry over numbers are distinguished in Nemeth code by using the notation "7777777". Any number above this row is recognized as a carry.

The restrictions become more complex when dealing with algebraic arrangements. In the following Nemeth code example:

```
 2x^2"-25x+ 7
25x^2"    +25
33333333333333
 27x^2"-19x+32
```

notice that The `x^2` terms, `x` terms and the constants are vertically aligned. The `+/-` sign are also vertically aligned.

Handling other spatial structures can be likewise complex. For example, there are at least 12 different symbols for spatially representing long division [4]. The division notations in Nemeth is formed using curved division signs made of either '(' or ')' on the right or left, horizontal separation lines which are either in the line above or below, vertical separation lines which are to the right or left, or slanting lines to the right or left. Since any division sign is composed of at least two different types of above mentioned symbols, Winsight has to scan two lines and analyze them before concluding that the input is indeed a division symbol.

In general, Braille to LaTeX translation of spatially arranged arithmetic/algebra cannot be done on a character by character or line by line basis; multiple lines need to be analyzed before conclusions regarding the type of spatial arrangement can be drawn. The same is true for spatial arrangements involving algebraic expressions (e.g., division of a polynomial by another polynomial), these have to be handled by separating algebraic expressions from spatial information, processing them separately, and combining them to create final output. The Winsight system handles all types of spatially arranged structures in Nemeth code.

3 Overview of the Winsight System

The Winsight system backtranslates not just mathematics but complete *mathematical documents*, i.e., documents that contain Mathematics mixed with text. Mathematics is seldom written by itself, it is generally embedded in text. In a Braille document, this text is written in contracted Braille. The Winsight system takes a hard copy Braille document as input, and produces a print file (in pdf) containing the equivalent Math and text. The composition of the Winsight system is shown in Figure 1. Arrows with the label of "M" indicate points at which the user has the option of examining the output produced by that module and changing it.

Winsight OBR Module: The Winsight system takes a JPG file containing image of a hard copy Braille sheet. This image is obtained by scanning the Braille sheet on an ordinary scanner. The scan records the image of shadows of

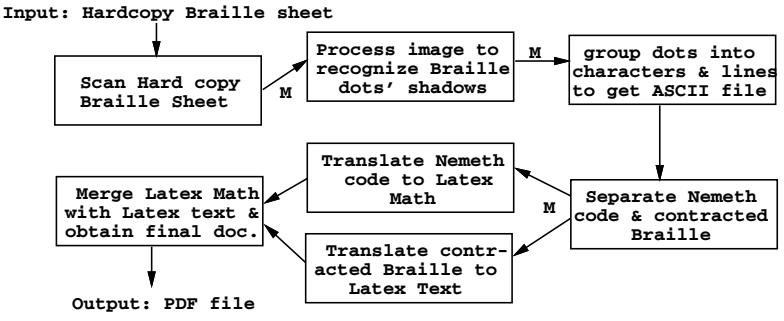


Fig. 1. The Winsight System Flow Diagram

the Braille dots. The user can edit the JPG file using the MS Paint program from within the Winsight system. This editing may be needed to remove smudges on the page or marks left after removing staples. Next, the Winsight system processes the image of the shadows to detect the Braille dots as "blobs."

Next, a sophisticated reasoning algorithm deduces the characters represented by these dots to create an ASCII Braille file. Numerous challenges arise in this step: since the Braille sheet is more likely to be manually produced, the spacing between the lines may be uneven (the lines may be too close), or the paper may be tilted during typing, causing one line of Braille to be slanted w.r.t. another. The paper may have been slightly moved in the middle of the line while typing, causing the remaining part of the line to be slightly misaligned. These problems arise more in hand-typed Braille document; Braille documents produced using an embosser are evenly spaced and do not have such errors. To the best of our knowledge, OBR systems available in the market are designed for documents produced on embossers, and thus are not suitable for hand-typed Braille sheets. In contrast, the Winsight system's OBR module works very well with hand-typed Braille sheets. However, on occasion it will ask the user a few questions (for example, if every Braille character present in the sheet consists of only dots present in the bottom two rows, then the system will ask the user if the top line contains only bottom two rows or top two rows, etc.) and issue warnings (e.g., if two lines are too close). The user can further change the output of the image recognition in two ways:

1. The user can examine the "blob" file and fix any image recognition errors by comparing it with the hard copy Braille sheet. This is helpful when the Braille sheet contains Braille characters that have been erased by back-spacing. In such cases, the erased dots don't completely go away, and their shadow may appear in the scanned image. The user does not need to have any knowledge of Braille for making these fixes.
2. The user can look at the generated ASCII Braille file, and fix its contents. This, however, requires knowledge of mapping between Braille characters and ASCII codes.

Winsight Separation Module: The Winsight system next attempts to automatically separate contracted Braille (Grade II Braille) text from Nemeth Braille Math code. This is a tricky process as it is hard to tell contracted Braille apart from Nemeth code. The same Braille character string can mean different things depending on whether we interpret it as contracted Braille or as Nemeth code [5]. For example, consider the Braille string `y "k t`. If we interpret it as Nemeth code, it will mean $y < t$, while if we interpret it as contracted Braille then it will stand for *you know that*. Likewise, the Braille string `h;e` means h_e in Nemeth code, but *hence* in contracted Braille, and `r+s` means `r+s` in Nemeth code, but *rings* in contracted Braille.

Thus, the general problem of automatically separating Nemeth Braille code from contracted Braille is quite difficult as one has to examine the context. The Winsight system uses a number of heuristics, including knowledge of English words as well as knowledge of various operators and mathematical patterns that can arise in Nemeth code to automatically annotate Nemeth Math expressions in the document. This annotation is done by enclosing Nemeth Math expressions within `;;` and `;'`, as stipulated by Dr. Nemeth [5]. The notation proposed by Dr. Nemeth recognizes a number of subcases, however, Winsight does not support those subcases. This is due to our belief that these subcases are hard to remember for students, and they unnecessarily complicate the notational convention. Future versions of Winsight may support it.

Winsight provides the user the option of viewing the annotated file. A separate option within Winsight allows the user to insert or delete the Nemeth math indicator tags. The math indicator tags, while internally stored as `;;` and `;'`, are displayed to the user as `<MATH>` and `</MATH>` for ease of visibility and modification. Note that the students themselves can always enclose the Math portion within the `;;` and `;'` tags while typing the Braille document on the Brailler. Expressions within such tags will be treated as Math.

Backtranslation: Once the Nemeth Math code portions and contracted Braille portions of the document have been separated, they are separately backtranslated.

1. The Nemeth code backtranslator in turn then separates mathematical expressions from spatial structures. Thus, given a matrix of of algebraic expressions, the individual algebraic expressions will be isolated and backtranslated separately, the structure of the matrix analyzed separately, and then the two put together to produce a LaTeX rendering of the Matrix. If the Nemeth code backtranslator fails to parse an expression, it is output unchanged. This failure in parsing is usually due to grammatical errors committed during typing the Braille document. The user can fix these errors in the annotated ASCII Braille file and backtranslate again. The Winsight system attempts to fix some errors automatically, but as is well known, automatic error detection and correction during parsing is extremely difficult. Note that our intent is that the backtranslation tool will be used by teachers of blind students. Generally speaking, the Winsight system will be used for grading answers to homework and exams, and in such cases, the teachers have a very good idea

as to what is expected as an answer, and so are likely to be able to fix the typos themselves. They can also seek the students' help in fixing the mistake, as the teachers themselves may not know Braille. Our practical experience indicates that with some rudimentary knowledge of Nemeth code, such typos are relatively easy to fix. To aid in this process, the Winsight system also provides an interface in which a user can type stand-alone ASCII Braille strings, and backtranslate them assuming either that it represents Nemeth code or that it represents contracted Braille. This facility allows a person who does not know the Nemeth code well to try the various fixes that could be possible for the typo made by the student.

2. The contracted Braille backtranslator examines each word, and using the standard mapping of English phonemes and syllables to Braille symbols, backtranslates the whole text. The contracted Braille notation also has context sensitivity and ambiguity built-in and many problems arise during the backtranslation. For example, the Braille character '4' can mean different things in different context. Thus, '4connect' in contracted Braille means *disconnect*, '4t4' means *dist.* (distance or district) while 'ha4ock' means *haddock*. Similarly, the number '0' in contracted Braille can be mapped to *was* or *by* if it occurs by itself. The Winsight system will output *by/was* for a stand-alone 0, relying on the sighted reader to figure out which one was intended.

Final Output: Once the Nemeth Math part and the contracted Braille parts are backtranslated, they are merged to produce the final document in LaTeX. The final LaTeX document is processed and displayed in PDF format, that a sighted user can then read. Note that the Winsight system has been extensively tested on real homework solutions. Below we show part of an example annotated ASCII Braille file produced from a scan (from an actual student homework), and the result of backtranslation.

Annotated ASCII Braille file

```
,homework    ,n ,y>i*
#a4 ;;  ?1/2#+?1/6# .k ?4/6#  ;'
#b4 ;;  ?3/4#+?5/8# .k ?11/8#  ;'
#c4 ;;     160  r2
          33333333333
         4o674
           4
          33333333333
           27
           24
          33333333333
            34
            32
          33333333333
             2 ;'
```

Resulting Print Output

Homework N Yarich

1. $\frac{1}{2} + \frac{1}{6} = \frac{4}{6}$
2. $\frac{3}{4} + \frac{5}{8} = \frac{11}{8}$
3. $4\overline{)674}$ $160\ r2$

$$\begin{array}{r} 4 \\ \hline 27 \\ 24 \\ \hline 34 \\ 32 \\ \hline 2 \end{array}$$

4 Conclusion

In this paper, we gave an overview of the Winsight system for backtranslating complete mathematical documents written using Nemeth code and contracted Braille to print. The system takes as input the scanned image of a Braille sheet, performs image processing to recognize Braille characters, and then backtranslates the resulting ASCII Braille file to LaTeX. The Winsight system is an innovative system, as there is no other system at present that can provide complete backtranslation. Future work includes refining our system to provide better error detection and correction, so as to minimize human intervention.

Acknowledgment. Thanks to Susan Osterhaus of TSBVI, Susan Jolly of Los Alamos, and Dipendra Manocha of the National Association for the Blind, New Delhi, India, for discussion and comments, and to the Department of Education and the National Science foundation for support. Views expressed in the paper are solely those of the authors.

References

1. Aho, A., Ullman, J., Sethi, R.: Principles of Compiler Design. Addison-Wesley, Reading, MA (1986)
2. Annamalai, N., Gopal, D., Gupta, G., Karshmer, A., Guo, H., INSIGHT,: A Comprehensive System for Translating Braille based Mathematical Documents to LaTeX. In: Proc. HCI'03. pp. 1245–1249 (2003)
3. Karshmer, A., Gupta, G., Geiger, S., Weaver, C.: A Framework for Translation of Braille Nemeth Math to Latex. In: Proc. ACM ASSETS'98, pp. 136–143 (1998)
4. Nemeth, A.: The Nemeth Braille Code For Mathematics And Science Notation 1972 Revision. American Printing House For The Blind (1972)
5. Nemeth, A.: The Nemeth Uniform Braille System (NUBS). Draft (April 2003) http://www.braille2000.com/brl2000/docs/NUBSdraft1.pdf
6. Scadden, L.: Making Mathematics and Science Accessible to Blind Students Through Technology. In: Proceedings of RESNA'96 (1996)
7. Lamport, L.: LaTeX A Document Preparation System. Addison Wesley, Reading, MA (1994)
8. http://www.mackichan.com

Authoring Tools for Structuring Text-Based Activities

Maria Grigoriadou and Grammatiki Tsaganou

University of Athens, Dept. of Informatics & Telecommunications
Ilisia Panepisthmiopolis GR-15784, Athens, Greece
gregor@di.uoa.gr, gram@di.uoa.gr

Abstract. In this paper we present research results and discuss recent directions concerning the structural analysis of science texts and cognitive aspects of text elements, such as causal relationships between text elements. We outline the process of structuring text material for the design of dialogue activities for science text comprehension in the educational environment of ReTuDiS (Reflective Tutorial Dialogue System). The system supports text comprehension using reflective dialogue activities, adapted to learners of different levels and its authoring tool, ReTuDiS*Auth,* supports users in authoring dialogue activities.

Keywords: Text structure, authoring tools, dialogue activities, cognitive categories and causal relationships.

1 Introduction

In approaching text comprehension, researchers examine issues that focus on assisting comprehension through the design of the text form [8]. Assuming that readers build mental representations of information contained in the text during the comprehension process, primary role should be attributed to the understanding of cognitive categories such as state, event and action and their causal and temporal relationships [2].

The design of the structure of text-based activities is important in order to enhance learning in an educational system. In order to make the information in such activities available to target users (students, teachers, researchers, authors, educators) new efforts have emerged to bring together novel methodologies and technologies. Authoring such activities demands an authoring system which involves knowledge acquisition, design process and managing a large amount of complex information [18]. Authoring tools offer the appropriate structure and guide authors to import and elaborate educational material (text, questions, dialogues etc.) [1].

Researchers have been investigated Intelligent Tutoring Systems (ITS) authoring tools almost since the beginning of ITS research and authoring systems have been built [16] [20]. An authoring tool is a generalized framework along with a user interface that allows non programmers to formalize their knowledge [16] [19] [24]. Part of authoring an ITS is the systematic decomposition of the subject matter into a set of related text elements. Each authoring system provides tools or cues which assist the author in this process of breaking down and elaborating the content to the necessary level of detail according to an instructional model. There are intelligent adaptive hypermedia systems like CALAT [18] and GETMAS [24] that their functions overlap those from

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 319–328, 2007.
© Springer-Verlag Berlin Heidelberg 2007

both the above categories. There are also expert systems, like Dempndtr8 [3], IRIS-tutor [2] which include rule-based cognitive models of problem solving expertise and observe learner behaviour in order to build a learner model.

In this paper we present the process of structuring text material for the design of dialogue activities for text comprehension in the environment of ReTuDiSAuth, which is the authoring tool of ReTuDiS system [22] [10]. In the second section, we outline theories behind text comprehension concerning cognitive relationships. In the third section, we comment about authoring tools and authoring process. In the fourth section, we present the main components of ReTuDiS and ReTuDiSAuth and focus on describing knowledge acquisition methods and authoring techniques offered to the authors in order to build dialogue activities for text comprehension. In the fifth section, we describe text structuring in the environment of ReTuDiSAuth. In the sixth section, we present evaluation results. Finally, we conclude and give our future perspectives.

2 Text Comprehension

Many studies on text comprehension have focused their interest on the sentence structure presented by the text [5] [15]. Sentence structure of a text could be organized on the basis of hierarchy in order to allow the importance of sentences in the text to be revealed [23]. The cognitive psychological approach in text comprehension suggests that the internal variables of the reader hold a primary role in text comprehension, such as his personal goals, interests and pre-existing knowledge. However, cognitive science does not ignore the influence of the text form, in which factors such as text cohesion and logical coherence of facts presented have been proved to be significant elements that facilitate its comprehension [7].

In order to examine the representation constructed by learners during the comprehension process of a text, primary role should be attributed to the understanding of the cognitive categories *state*, *event* and *action* [4]. The term *state* is static and describes a situation in which no change occurs in the course of time. The term *event* refers to an effect, which causes changes but is not provoked by human intervention. The *event* can be coincidental or provoked by human intervention, e.g. by a machine. An *action* causes changes but is originating by a man. This consideration deals with text comprehension as the attribution of meanings to causal connections between occurrences in the text. Learners compose a representation of the text, which contains the cognitive categories: event, state and action. For the interpretation of learners' cognitive processes their discourse is analysed, in order to trace the recognition (or not) of the three cognitive categories.

Furthermore, text analysis in relation to the cognitive categories does not suffice [4]. The organization and structure of cognitive representation should be examined on micro and macro-levels. Mental representations capture elements of the surface text, of the referential meaning of the text, and of the interpretation of the referential meaning, thus constructing a micro-world of characters, objects, spatial settings, actions, events, feelings etc. [15].

The person who reads a text gradually constructs the microstructure of the text representation, i.e. the states, event and compound actions of the world described in

the text as well as the time and causal relationships that interlock those structures. In order a person, to be able to explain the operation of a technical system on a micro-level scale, has to construct a representation of the "natural flow of things", where every new event should be causally explained by the conditions of events which have already occurred [4]. The creation of a text that allows a precise description of a technical system and facilitates readers in constructing its microstructure representation must involve: (a) the description of the units that constitute the system based on the causal relationship which unites them and (b) the description of event sequence taking place in these units in respect of the cause affecting them as well as of the changes they bring to the state of the system.

On macro-level, the development of the macrostructure by readers is achieved through the reconstruction of the microstructure and the establishment of a hierarchical structure with goals and sub-goals. The creation of a text which facilitates readers in constructing its macrostructure representation for a system must involve the teleological hierarchical structure of goals and sub-goals of the various operations as well as their implications.

For example, in a technical text describing a local network, the network constitute a system of units (server, workstations and peripherals) connected according to a topology [10]. Messages are sent and received between the local network devices causing events. There are causal and temporal relationships between events [13]. At the end of a sequence of events the system changes its state returning from one state to another.

3 Authoring Tools

ITS authoring is both a design process and a process of knowledge articulation. While authoring tools are becoming more common and proving to be increasingly effective they are difficult and expensive to build. Authoring tools use methods to achieve the following goals [1] [18]: a) decrease the effort of authoring (time and cost), b) allow others to take part in the design process c) help the author articulate or organize his domain knowledge d) support good design principles concerning the pedagogy and the interface and e) allow quick evaluation cycles.

Authoring tools achieve the above goals using a number of methods. Authoring systems use methods to simplify and automate authoring and knowledge acquisition. Part of authoring an ITS is the systematic decomposition of the subject matter into a set of related elements, for example a hierarchy. Each authoring system provides tools or cues which assist the author in this process of breaking down and elaborating the content to the necessary level of detail according to an instructional model.

Authoring tools allow non-programmers to build tutors by incorporating a particular model or framework to scaffold the task [18]. Learner modelling process requires making certain choices, and it is in these choices that the learning process is located [14]. We do not learn much from looking at a model, we learn from models by building them and using them [17] [12]. Learning from building models involves finding out what elements fit together in order to represent the world of the model. The design of dialogue activities for adaptive learning supported by appropriate authoring tools attracts the interest of many researchers and educators in inventing

new methodologies for effective teaching and learning. The authoring process activates authors to decompose the subject matter into a set of related elements to discover what elements fit together in order to represent a concept, for example a hierarchy. The authoring process, as a process of choosing, organizing, structuring and linking educational material becomes a process of learning.

Authoring tools for text comprehension have to discover and offer mechanisms which help authors design activities for the diagnosis of learners' difficulties in comprehending texts. They offer the appropriate structure and guidance in order the author to be able to import and elaborate educational material (text, questions, dialogues etc.). There has been a growing concern about scientific text comprehension [5]. Efficient teaching and learning requires that educators should be familiar with the difficulties which learners are likely to face.

4 ReTuDiS and ReTuDiS*Auth*

4.1 ReTuDiS System

ReTuDiS is a diagnosis and open learner modelling tutorial dialogue system for text comprehension. The system infers learners' cognitive profile in order to construct and revise the learner model with the learners' participation [21]. ReTuDiS consists of two parts: the Diagnosis part and the Dialogue part.

The diagnosis part of ReTuDiS approaches learner's text comprehension supporting the theory of Baudet & Denhière that learner's representation of the text contains the cognitive categories: event, state and action [4]. The system engages learners in an activity which includes reading comprehension of text and answering question-pairs by selecting between given alternative answers. Learners' answers are used for diagnosing learners' text comprehension. Learners have to study all the text to comprehend it, to compare each factor with the others and then select answers from the given alternative answers, in order to express their position on certain issues and support it by a justification. The diagnosis part infers learners' cognitive profile and his learner model.

The underlying theory beyond the dialogue part of ReTuDiS is the Theory of Inquiry Teaching [6]. ReTuDiS approaches dialogue activities based on theories of dialogue management, strategies, tactics and plans which promote reflection in learning. The dialogue part is based on the learners' cognitive profile, inferred by the diagnosis part, the learners' answers to question-pairs and the selected dialogue strategy offered by the system. The dialogue part of ReTuDiS engages the learner in personalized reflective dialogues in order to revise the learner model with the participation of the learner [22]. The dialogue generator activates the appropriate for the learner sequence of dialogue-parts, and using the dialogue plan, dynamically constructs the individualized learning dialogue.

4.2 ReTuDiSAuth

ReTuDiS*Auth*, the authoring tool of system ReTuDiS, offers an environment that lays out the appropriate parameters an author needs to define. The authoring tool supports users registered as teachers or administrators. Teachers have the authority to create

new activities or edit existing ones. Administrators of the system have the authority to manage the base of the users of the ReTuDiS and the educational material. The environment offers the *tools* and the *shell* (Figure 1). The tools, which add interactivity to the system and support authors to import educational material, are the following:

Text fields. These fields are designed to help authors enhance their own educational material into the system for example, titles of activities, texts, questions.
Pop-up menus. The menus are designed to help authors select from predefined by the teacher or by the system values or defaults such as: categories of activities, characterizations of answers, teaching strategies.
The knowledge base. This data base includes the whole educational material.
Association buttons. They are buttons designed to help authors establish causal relationships between text elements such as factors of a text and make associations such as between educational material and learners' profiles, teaching strategies and learners' profiles, learners' profiles and dialogue plans.
Guidance tips. They are information tips designed to support the author by giving back the appropriate feedback to his actions that is confirmation or not of the completion of each step.
Administrative tools. These tools are for managing the lists of users, the roles of the users (teachers or students), the categories of activities, the activities, reports on carried out activities (log files for each student).

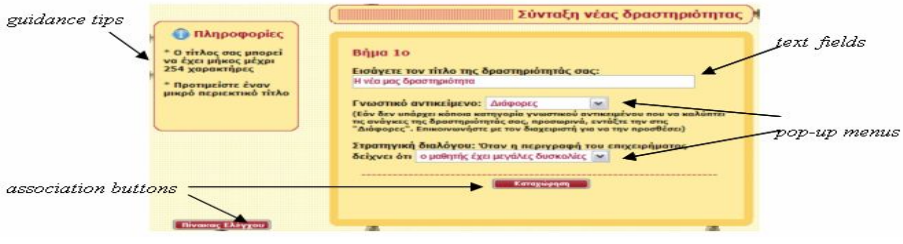


Fig. 1. The ReTuDiSAuth: Tools for structuring and elaborating a dialogue activity

The shell delivers the educational material according to the instructions generated by the author using the tools in combination with its predetermined defaults. A semantic network is created by the author using the association buttons. The nodes of the network represent the elements of the material and the branches represent possible paths followed by the learner while participating in instructional activities. While delivering the educational material to a learner in a specified manner the shell constructs his learner model. The values of the learner model change over the activity course as the learner participates in dialogues.

5 Text Structuring with RETUDIS*Auth*

The structuring process needs the description of the domain knowledge according to the system's specifications. ReTuDiS*Auth* involves the author in the following

processes concerning text structure, questions with alternative answers and dialogue structure:

5.1 Text structure

Structuring the text by separating it in paragraphs called factors. The author focuses on assisting comprehension through the design of the text form [22]. He forms the text in paragraphs, called factors, each of which represents a cognitive category *state, event* or *action.*

Titling paragraphs. The author gives a phrase as a title for every factor. The title helps in organizing the structure of the activity.

Specifying the cognitive categories. The author specifies the cognitive categories involved and the number of them. For example, in case of historical text are used three cognitive categories: state, event and action [21] [11].

5.2 Questions and Alternative Answers

Forming questions with alternative answers. For every titled factor the author submits a question-pair to the system and its related alternative answers. The first question in the question-pair is related to this factor and the learner's alternative answer concerning this question is declared to the system as *position*. The second question is related to the learner's justification concerning a position and is declared to the system as *justification*. Position and justification represent the causal relationships in the text.

Forming classifications. The author classifies the alternative answers concerning position and justification as *scientific*, *towards- scientific* or *non- scientific*. For every question-pair the combination of the learner's *position* and the corresponding *justification* constitutes the learner's *argument*. Arguments are classified as *complete,* when both position and justification are *scientific*. Otherwise the argument is *non-complete*. The author defines the different degrees of *argument completeness*. Possible values of argument completeness are: c*omplete, almost complete, intermediate, nearly incomplete* and *incomplete.*

5.3 Dialogue Structure

Forming dialogues. The author creates a library consisting of *specific* dialogue-parts for all combinations of possible answers and associates them with the corresponding answers. Each *specific* dialogue-part is designed to remedy a particular learning difficulty. The specific dialogue-parts are dependent on the specific text. The *specific* dialogue-parts of different types are associated with predefined and embedded in the system dialogue tactics

Forming dialogue tactics. Dialogue tactics, inspired by the general teaching strategies of Collins [6] [8], are hints or Socratic-style dialogues. Tactics correspond to different levels of dialogue concerning the specific subject matter and involve learners in activities which promote reflection. The author defines the dialogue tactics which have the following forms: (a) picks positive or negative examples, (b) picks

counterexamples, (c) generates hypothesis, (d) makes learner to form hypothesis, (e) makes learner to test hypothesis, (f) entraps the learner, (g) traces consequences to a contradiction or faulty knowledge of a learner and (h) promotes questioning authority.

Selecting dialogue strategy. The choice of the dialogue strategy is decided in the beginning. Example of a strategy embedded in the system is the following [11]: "The system sorts learners' argument classifications in a list according to decreasing degree of argument completeness. The tutorial dialogue begins with a discussion about the factor for which the learner seems to face less learning difficulties. The system generates the sequence of dialogue-parts for this factor. Then the system prepares the next dialogue-part, based on the results of the previous dialogue-part".

Selecting dialogue tactics. Predefined dialogue tactics are accessed throughout a pop-up menu. The author selects a predefined dialogue tactic and formulates the dialogue-part.

Planning dialogue. For the selected teaching strategy and depending on the learner profile the system constructs the initial dialogue plan for the learner. The system uses: (a) the *general* dialogue-parts, which include typical dialogue-parts, concerning participation of the learner in dialogue, encouragement, motivation, agreement or not with the system, guidance etc. and (b) the *specific* dialogue-parts that were previously entered by the author according to the appropriate dialogue goals and tactics. During entering the author has made the appropriate associations between contradicting answers (contradictions between learner's position and his justification concerning causal relationships in the text) and dialogue-parts for all possible combination of answers. So as, the system becomes able to initiate the dialogue and generate dynamically the appropriate dialogue in response to the learners' feedback during the dialogue process.

Defining learners' profiles. Learners can be described as belonging to one of a set of author-defined learner profiles taking into account the number of learner's arguments with high degree of argument completeness.

6 Evaluation

Formative evaluation, concerning the use of ReTuDiSAuth for text structuring, was conducted with the participation of 26 postgraduate students and 6 experts in informatics and history domains at the University of Athens. Evaluation aimed at further revisions, modifications and improvements of the system [18]. The participants were given explanations about the aims of the authoring tool.

Students were asked to participate in the evaluation process and perform representative tasks: (a) to prepare source material of their choice (b) to use the system for the construction of dialogue activities. Each student proposed a two pages text, three question- pairs with alternative answers involving causal relationships and specific dialogues-parts (Figure 2).

Experts used the material proposed by the students in order to identify and comment issues concerning specific problems or deficiencies users face in formulating learning goals, questions and tutoring dialogues and the educational benefits of the process.

Local Network Operation

…..A bus topology is designed with each node connected directly to a high-data speed bus. Nodes communicate across the network by passing packets of data through the bus (they read and write data -in the form of packets). Packets placed on the bus, transfer messages to nodes. A message includes the receiver's address, which specifies the network address of the target node. A node watches the bus continuously and reads the target address of each packet. After that, the node compares the address with its own, and if they are the same, then reads the message of the packet, otherwise ignores it. When a node is ready to broadcast a message, waits until the bus is free and then begins passing it to the bus. If a node uses the bus it watches it and can be aware of any other node using the bus at the same time. In that case both nodes stop using the bus waiting until one of them accidentally attempts to us it. When a limited number of packets are simultaneously transmitted throughout the bus, them this competence strategy is successful. The bus topology network can work even in case of disconnection of a node…..

Question-pair 3

3a) In a bus topology network, what happens in case there is an interruption (a cut off) of the bus. Select one of the following answers.

1. the network crashes (non scientific)
2. the network continuous to work properly (towards scientific)
3. the network is divided into two independent networks each one working properly (scientific)

3b) Justify your answer by selecting one of the following answers.

1. because all nodes are connected with the bus and they cannot communicate if there is a cut off (non scientific)
2. because all nodes have spare connections between each other that can operate without the bus (towards scientific)
3. because a bus network needs only a central bus to connect the nodes to (scientific)

Fig. 2. Text fragment and questions with alternative answers given by a student

Both students and experts were given a questionnaire and commented about usability, learnability and efficiency:

- the depth to which the system can infer a learner's knowledge, respond accordingly and teach
- if the system can support dialogue activities on different knowledge domains
- how easy non-programmers can learn to use the system
- how quickly a trained user can build dialogue activities
- the amount of resources needed to produce dialogue activities

Moreover, experts were asked to comment about:

- how much the underlying instructional model of the system constrains the author
- the sources of teaching and domain expertise
- the level of expertise /background of the target authors.

In general, most of the experts faced minor difficulties in using the interface. Experts spent more time to overcome difficulties in structuring the text and matching text paragraphs with cognitive categories. Experts commended about the quantity and the quality of questions made by the students. They identified as beneficial the method used for training students, which may be potential teachers, for the improvement of their authoring skills for text- based dialogue activities.

7 Conclusions

In this work we presented research results of the educational benefits after the use of ReTuDiS*Auth* as an authoring tool for structuring educational text material. Students experimented in the environment and designed text-based dialogue activities that promote learners' reflection. Evaluation got hold of representative users: graduate students and experts. Results of the current study indicate that authoring makes students improve their authoring skills and become familiar with text structuring. Experts found the system appropriate for the education of postgraduate students in reflective dialogue environments.

In our future plans we foresee improvement of the system concerning direct specification of causal connections between text elements during text structuring. Moreover, we plan further research into the evaluation of the system in classroom conditions and compare results in different knowledge domains.

References

1. Ainsworth, S., Major, N., Grimshaw, S., Hays, M., Underwood, J., Williams, B., et al.: REDEEM: Simple Intelligent Tutoring Systems from Usable Tools. In: Murray, T., Blessing, S., Ainsworth, S. (eds.) Authoring Tools for Advanced Technology Learning Environments, pp. 205–232. Kluwer Academic Publishers, The Netherlands (2003)
2. Arruarte, A., Ferrero, B., Fernandez-Castro, I., Urretavizcya, M., Alvarez, A., Greer, J.: The IRIS Authoring Tool. In: Murray, T., Blessing, S., Ainsworth, S. (eds.) Authoring Tools for Advanced Technology Learning Environments, pp. 233–267. Kluwer Academic Publishers, The Netherlands (2003)
3. Blessing, S.: A Programming by Demonstration Authoring Tool for Model-Tracing tutors. In: Murray, T., Blessing, S., Ainsworth, S. (eds.) Authoring Tools for Advanced Technology Learning Environments, pp. 93–119. Kluwer Academic Publishers, The Netherlands (2003)
4. Baudet, S., Denhière, G.: Lecture Comprehension de Texte et Science Cognitive, Presses Universitaires de France, Paris (1992)
5. Brown, A.L., Day, J.D.: Macrorules for summarizing texts: The development of expertise. Journal of Verbal Learning and Verbal Behavior 22, 1–14 (1983)
6. Collins, AI. A: Sample Dialogue Based on a Theory of Inquiry Teaching. In: Reigeluth, Ch. (ed.) Instructional Theories in Action, pp. 181–199. Lawrence Erlbaum Associates Inc. Hillsdale (1987)
7. DeCorte, E., Verschaffel, L., DeWin, L.: Influence of rewording verbal problems on children's problem representations and solutions. Journal of Educational Psychology 77, 460–470 (1982)

8. Graesser, A.: Teaching Tactics and Dialog in Auto-Tutor. International Journal of Artificial Intelligence In Education 12, 257–279 (2001)
9. Graesser, A., Tipping, P.: Understanding Texts. In: Bechtel, W., Graham, G. (eds.) A Companion to Cognitive Science, Blackwell, Malden MA (1999)
10. Grigoriadou, M., Kanidis, V.: Cognitive Aspects in Teaching the Computer Cache Memory with Learning Activities based on a Coherent Technical Text and a Simulation Program. In: Proceedings of the 6th Hellenic European Conference on Computer Mathematics & its Applications (HERCMA03) - Minisymposium: Informatics in Cognitive Sciences, pp. 235–429. LEA Publishers, Athens, Greece (2003)
11. Grigoriadou, M., Tsaganou, G., Cavoura, T.: Historical Text Comprehension Reflective Tutorial Dialogue System. Educational Technology & Society Journal, Special issue 8(40), 31–40 (2005)
12. Jonassen, D.: Model Building for Conceptual Change: Using Computers as Cognitive Tools. In: Grigoriadou, M., Raptis, A., Vosniadou, S., Kynigos, X (eds.) Proceedings of the 4rd Panellenic Conference with International Participation: Information and Communication Technologies in Education (ETPE2004) Athens, Greece pp. 4–17 (2004)
13. Leon, J., Penalba, G.: Understanding Causality and Temporal Sequence in Scientific Discourse. In: Otero, J., Leon, J., Graesser, A. (eds.) The Psychology of Science Text Comprehension, Lawrence Earlbaum Associates, Publishers, London (2002)
14. Kay, J.: Learner control. User Modeling and User.-Adapted Interaction 11, 11–127 (2001)
15. Kintsch, W.: Compréhension: a paradigm for cognition. Cambridge University Press, UK (1998)
16. Koedinger, K., Anderson, J.: Intelligent Tutoring Goes to the Big City. International Journal of Artificial Intelligence in Education 8, 30–43 (1995)
17. Morgan, M.S.: Learning from Models. In: Morgan, M.S., Morrison, M. (eds.) Models as mediators: Perspectives on natural and social science, pp. 347–388. Cambridge University Press, Cambridge (1999)
18. Muray, T.: An Overview of Intelligent Tutoring System Authoring Tools: Updated analysis of the state of the art. In: Murray, T., Blessing, S., Ainsworth, S. (eds.) Authoring Tools for Advanced Technology Learning Environments, pp. 491–544. Kluwer Academic Publishers, The Netherlands (2003)
19. Ritter, S., Blessing, S.: Authoring tools for Component-Based Learning Environments. Journal of the Learning Science 7(1), 107–132 (1998)
20. Schultz, K., Bratt, E.O., Clark, B., Peters, S., Ponbarry, H., Treeratpituk, P.: A Scalable, Reusable, Conversational Tutor: SCoT. In: Proceedings of the 11th International Conference on Artificial Intelligence in Education Workshop: Tutorial Dialogue Systems, Sydney, Australia, pp. 367–377 (2003)
21. Tsaganou, G., Grigoriadou, M., Cavoura, Th.: W-ReTuDiS: a Reflective Tutorial Dialogue System. In: Grigoriadou, M., Raptis, A., Vosniadou, S., Kynigos, X. (eds.) Proceedings of the 4rd Panellenic Conference with International Participation: Information and Communication Technologies in Education, Athens, Greece, pp. 738–746 (2004)
22. Tsaganou, G., Grigoriadou, M., Cavoura, Th., Koutra, D.: Evaluating an Intelligent Diagnosis System of Historical Text Comprehension. Expert Systems with Applications 25(4), 493–502 (2003)
23. Van Dijk, T.A., Kintsch, W.: Strategies of discourse comprehension. Academic Press, New York (1983)
24. Wong, W.K., Chan, T.W.: A Multimedia Authoring System for Crafting Topic Hierarchy, Learning Strategies and Intelligent Models. International Journal of Artificial Intelligence in Education 8(1), 71–96 (1997)

VoxBoox: A System for Automatic Generation of Interactive Talking Books

Aanchal Jain and Gopal Gupta

Department of Computer Science
University of Texas at Dallas, Richardson, TX 75080

Abstract. We present the VoxBoox system, a system for making digital books accessible to visually impaired individuals via audio and voice. This is accomplished by automatically translating a book published in HTML to VoiceXML, and then further enhancing this VoiceXML rendering of the book to enable listener-controlled dynamic aural navigation. The VoxBoox system has the following salient features: (i) it leverages existing infrastructure since the book that is to be made accessible need only be published digitally using HTML on the visual Web, (ii) it is based on accepted Web standards of HTML and VoiceXML and thus books can be made accessible inexpensively, and (iii) it is user-centered in that the listener (the user) has complete control over (aural) navigation of the book. In this paper, we present details of the technologies that make the VoxBoox system possible, as well as the details of the system itself. A prototype of the VoxBoox system is operational.

1 Introduction

Even though we have witnessed great leaps in technology in recent times, access to documents for people with print disabilities is still not seamless. Almost everything is digital today: books and papers are published online; even printed books are composed with the aid of the computer first and then they are printed in hard copies. However, even with books and papers available in digital form, there are significant impediments to making them accessible to the visually impaired. In this paper, we report on our efforts to make books easily accessible to the visually impaired readers by automatically turning them into talking books browsed using audio and voice. Our goal is to produce a solution that (i) leverages the existing infrastructure, (ii) uses widely accepted standards, (iii) uses existing inexpensive technology, and (iv) is user-centered. The culmination of our efforts is the VoxBoox system. The design of the VoxBoox system is based on our earlier efforts to make the Web accessible to the visually impaired individuals [7], a significantly harder problem if we observe the criteria mentioned.

The VoxBoox system automatically converts digital books, published on the visual Web using HTML, into talking books. The VoxBoox system does so by automatically translating books published using HTML markup to VoiceXML markup. HTML is the notation used for marking-up documents on the Web for

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 329–338, 2007.
© Springer-Verlag Berlin Heidelberg 2007

proper visual display. VoiceXML, in contrast, is a notation used for marking-up voice and audio documents. Thus, as soon as a book is published on the visual Web using HTML, with the help of our system it immediately becomes available in voice/audio. The VoxBoox system further enhances the VoiceXML rendering of the book to facilitate interactive, user-controlled voice and audio-based navigation of its content. Thus, using the VoxBoox system, not only can a (blind or sighted) individual hear the book through audio, he/she can also interactively navigate the book via voice commands.

The VoxBoox system leverages the current infrastructure of the Web to make books accessible to visually impaired. The VoxBoox system relies on the technologies of HTML and VoiceXML. HTML is a well-accepted standard for publishing information on the Web. A significant number of books are available in HTML format on the Web (see for example http://onlinebooks.library.upenn.edu), or they are available in formats that can be readily translated into HTML (such as pdf). Likewise, VoiceXML is a standard, well-established technology, albeit most of its usage has been in providing automated Voice based interaction over the phone [10]. With the help of VoiceXML technology, the VoxBoox system makes it possible for books to be interactively and aurally accessible over cell phones and land-line phones.

Our efforts to make published material available to the blind are along the lines of the DAISY (Digital Accessible Information SYstem) project [3]. The rest of this paper describes the VoxBoox system in detail. Finally, note that our Voice based approach not only enables eyes-free reading for the visually impaired, it also enables voice interactions with the document for the sighted individuals and thus is useful for sighted population as well.

2 Talking Books

Digital Talking Books can be regarded as electronic files, which are accessible via audio. According to DAISY [3], an organization dedicated to helping people with print disability, talking books are a collection of electronic books that are presented to visually impaired people through human or synthetic speech [1,2,4]. Currently, there are two common solutions to obtain talking books: recorded cassettes and Digital Talking Books (DTB) by the DAISY consortium.

2.1 Recorded Cassettes

Recorded cassettes are audio files that have specially been recorded on an audio cassette using human voice in order to help blind individuals "read." There are many libraries across the United States such as the California state library, Washington Talking Books and Braille library, Texas State Library, etc., which produce such cassettes. There are about 672,000 such recorded cassettes available around United States libraries [2]. Approximately 1000 copies each are produced annually by the National Library Service for the blind and physically Handicapped (NLS). These are distributed on 4-track 15/16 ips cassettes through 140

regional and subregional libraries [1]. The key point here is that every year there is a significant effort and money that is spent on preparing these cassettes. They are not only expensive but are also played and navigated manually. Also it is not possible to record millions of publications that are produced every year instantaneously. There are still so many books that are published and are not available to blind individuals because they have not been recorded yet.

2.2 Daisy Talking Books

Daisy (Digital Accessible Information System), formed in May 1996 by talking book libraries, is a consortium that works to provide transition from analog to digital talking books. The vision of the DAISY consortium is to make all published material accessible to people with print disabilities inexpensively [3]. The Digital Talking Books (DTB) by DAISY is a PC compatible software audio player. The player plays files that have been developed in accordance with the DTB format and provide facilities like book marking, highlighting, keyword search, etc [4].

The problem with the DAISY's approach is that, first, a book has to be published in the DTB electronic format, which can be a difficult task (however, it is possible for this conversion to be automated) and, second, navigation has to be done via a computer keyboard. The main disadvantage of the DAISY approach is that a computer is needed to use the system. In many developing countries such as India and China, people can afford mobile phones but they still cannot afford computers. Digital talking books available in the DAISY format can also be played using a DTB player (similar to a cassette player), however, this player has to be purchased separately.

3 The VoxBoox System

VoxBoox is a system developed by us to enable HTML-coded digital books to be made accessible to visually impaired individuals on the fly. The book can be navigated aurally (i.e., using audio and voice) using phones, PDA, mobile and even computers. With the VoxBoox system, blind as well as sighted individuals can easily navigate digitally published books via voice and audio. Because the system can be accessed over the phone, the system can be used to aurally browse books even at times when a computer is not available or cannot be operated (such as while driving).

3.1 Navigation of HTML and VoiceXML Documents

As mentioned earlier, we rely on VoiceXML [6,10] to make digitally published books accessible. VoiceXML is a W3C (World Wide Web Consortium) standard markup language for marking up documents that are to be played using audio and that receive input via Voice. VoiceXML is the audio/voice analog of HTML,

the mark-up language used for coding the documents on the visual Web. Just as there are visual browsers for visual reading of HTML documents, there are voice browsers, which interpret VoiceXML files for browsing with audio/voice. However, before we discuss how content can be made available in audio/voice via VoiceXML, we need to consider the following issues.

First, there are billions of HTML coded web pages available today on the visual Web. To make all these pages also available in VoiceXML would not only be expensive but also be very cumbersome, if proper methodology is not used. Indeed, companies have designed voice portals from scratch (such as those by Tell Me Studio; dial 1-800-TELLME to get a flavor) to make certain type of information available in audio/voice (stock quotes, directions, sports scores, etc). Voice portals have not proliferated due to the considerable expense involved in developing aurally navigable pages from scratch. A better (inexpensive) approach, that we have championed, is to automatically translate HTML coded pages into VoiceXML coded pages [8]. Following this path, the current Web infrastructure can be leveraged to make, at least theoretically, the entire Web content accessible [7] to visually impaired individuals.

Second, aural browsing is inherently sequential. This is in contrast to visual browsing where many things can be quickly scanned in one visual glance. Thus, it is of utmost importance to provide extensive navigational control (such as skipping text, moving backwards and forwards) to the user during aural browsing. Along with the inherent limitation of sequentiality, there are other limitations that are present in VoiceXML itself which significantly curtail users' navigational freedom during aural browsing. VoiceXML is a markup language, which has plain text included within "form" tags. A "goto" tag provides navigation from one form to another. Unfortunately, this kind of design leaves the control of how a VoiceXML page is navigated completely in the hands of the writer of the Web page [5]. The user has absolutely no say, apart from providing responses to questions asked during navigation. Given that the author of a VoiceXML page cannot guess all possible ways in which a user might interact with a page, the approach in which the author of the page tries to guess all possible navigation scenarios in advance does not work either. In order to avoid serial reading, which could be boring, as well as giving more flexibility to the listener, adequate navigation controls should be provided to the user. One-way to solve this problem is to enhance the page with additional control facilities during the translation of the document so that the user has a better browsing experience. This is the approach we take [9]. The VoiceXML page is enhanced so that voice commands such as skip (move to next form), back (move control to previous form), start (go to beginning of document), end (go to end of document), repeat (repeat the current form from the beginning) and pause (suspend reading until user says resume) will work during browsing. Additionally, to permit free form navigation, the listener can place speech bookmarks (we term them voice anchors) on various forms (paragraphs). Then by uttering these bookmarks later, users can move to an arbitrary paragraph in the document at their will. The standard

visual web allows user to bookmark certain URLs, but it does not allow them to bookmark a specific paragraph within the document (this can only be done by the page author using HTML anchors). We also allow users to bookmark the whole VoiceXML page, so that they can navigate back and forth between different pages just by uttering these bookmarks. Finally, keyword based search is also permitted: users can utter a keyword and navigation will jump to the form that contains that keyword. Note that in all cases, the user may have to spell the word the first time they speak it while placing anchors or bookmarks, or during keyword based search.

Note that the approach we take satisfies all the criteria under which we wanted to design our system:

- Because contents are obtained by translating HTML pages, existing infrastructure is leveraged;
- Because we obtain content by an automated translation, no effort is required to make a newly published book accessible. As soon as it is available digitally in HTML, it becomes available in VoiceXML as well.
- Because we rely on HTML and VoiceXML, widely accepted standards for the Web, no new language or convention needs to be learned
- Because we enhance the VoiceXML pages with additional control and voice anchors, listeners have complete control over their browsing experience.

3.2 Usage Scenario for the VoxBoox System

To understand the VoxBoox system, consider an example scenario. Let us say the user wants to read the book titled "Oliver Twist." The user will dial a toll free number. which will connect him/her to a Voicebrowser (a Voice browser is the VoiceXML counterpart of the visual HTML browser). The first page encountered by the user is an index page, which seeks input from the user. The user will speak the title of the book (Oliver Twist). Because of the limitations of general speech recognition, however, the inputs may have to be spelled the first time. This is very similar to retrieving books over the visual Web, except that the input is via voice and output is via audio. This query is transformed into a search over the regular Web, from where the requested book (published in HTML) is retrieved. The HTML coded text for the book is then automatically translated into the VoiceXML format by our system, which is then played to the user who can listen to the book. During translation, the generated VoiceXML page is automatically augmented with features to allow dynamic navigation of its contents through speech commands and voice anchors. Speech commands allow the user to move forward and backward, skip and repeat text, etc. They also allow the user to perform a keyword based search. Additionally, users can place speech labels as bookmarks (called voice anchors) on passages and return to them later. Using Voice anchors, a listener can mark important passages, and move back and forth between different passages, much like flipping pages of a physical book.

4 VoxBoox System Architecture

Next we present the architecture of the VoxBoox system. VoxBoox converts an HTML-published into VoiceXML and further enhances it to incorporate voice-based navigation for the user. This is diagrammatically illustrated in figure 1.

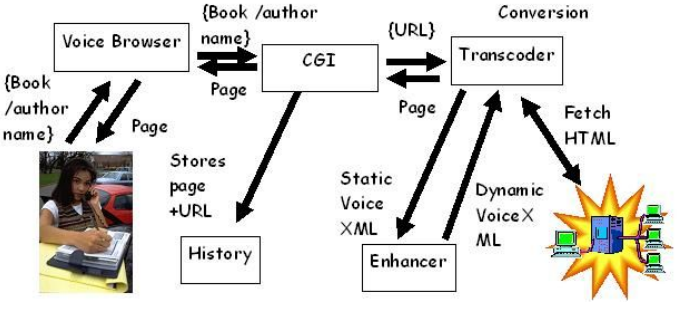


Fig. 1. Architecture of Interactive Talking Books

Before going into further details, let us first discuss the flow of control in the VoxBoox system. When the user dials a toll free number, he/she is connected to the Voice browser and the menu page is played (the names of the books available is read out one by one; the user can of course aurally navigate this menu page as well). The user utters the name of the book they want to access (there are other options available as well, such as searching by author, etc.); the browser then calls a CGI (common gateway interface) executable to handle the request. A URL is generated on the fly and is passed on to the transcoder (the module that performs the translation of HTML to VoiceXML), which fetches the HTML Web page of the corresponding book from the Web server (we use the Web server at onlinebooks.library.upenn.edu). The transcoder then converts the HTML page to VoiceXML and sends it to the enhancer (the enhancer adds control support for skip/pause/etc. as well as voice anchors to the VoiceXML page), which after enhancing sends it back to the transcoder, which sends it back to the CGI. The CGI sends the page to the browser and saves it in the history for further reference. The browser then plays the page to the user who can now navigate it aurally.

The structural design of VoxBoox can be broken down into the Voice browser, CGI (Common Gateway Interface), the transcoder, web server and enhancer. Figure 1 shows the pictorial representation of the interactions between these components. The Voice browser and the web server are independent plug-ins, which have already been developed. Voice browsers have been developed by companies such as IBM, Motorola, Tell Me Studios, etc. Our work thus mainly focuses on developing the CGI code, transcoder and the enhancer modules. To handle different type of user requests, different CGI modules have been developed. Whenever a new page is retrieved, a CGI executable is called which acts

as an interface between the transcoder and the Voice browser. Whenever voice anchors are placed and recalled, respective CGIs are called to place and recall these anchors. The transcoder module works as a translator. It converts an HTML page to static VoiceXML on the fly [7]. Whenever a HTML page is retrieved from the server, it is translated line by line to its VoiceXML version. To allow dynamic, interactive navigation of VoiceXML pages, the VoiceXML output by the transcoder is sent to the enhancer module. The enhancer inserts code and tags to allow users to place dynamic anchors in various paragraphs and retrieve them later on the fly. It also explicitly adds tags and code to allow users to skip through the paragraphs in the document. Users can repeat sections, go to the end or beginning of the document and can pause the document whenever they want. The user can even go back and forth between pages hyperlinked to each other, just as in the visual Web [7]. Next, we summarize the design and implementation of these three components.

Transcoder: As mentioned earlier, the transcoder is responsible for translating HTML pages to VoiceXML. It is implemented using Java by our group [8]. It has 2 phases: a parsing phase and a translation phase. Once it receives HTML pages from the Web Server in phase one, it parses the HTML file to obtain a parse tree. In the second phase, the parse tree is recursively translated to a semantically equivalent VoiceXML page. The transcoder semantically transforms every HTML tag to a corresponding VoiceXML tag. For this, the HTML pages has to be well-formed and should be free of extraneous information such as Java scripts [7]. To overcome this difficulty we first pass it through a filter, which removes all the unnecessary details from the HTML pages and also makes it well formed before passing it to the transcoder.

Enhancer: The foremost advantage of visual Web is to have multi directional reading. The user can skip the irrelevant material and read only the pertinent text. To provide such facility to the aural Web, the enhancer [9] adds certain tags and code to the static VoiceXML output by the transcoder. This enhancement will allow skip, back, start, end, pause, and repeat voice commands to work, along with voice anchors and keyword based search. The enhancement process is briefly described next.

The texts inside VoiceXML document are categorized into “form” tags. Each form has a unique form id. The flow of the document is determine by the “goto next =” tags. In static VoiceXML pages, the “goto next” tag is set to point to the next form tag, hence each form is played sequentially, and therefore the content of a VoiceXML page is heard serially. The writer of the document is responsible for incorporating navigation and giving users some control over the order in which forms are read. Our system incorporates the utility to skip, repeat, back, pause, resume, end and start in a document. These utilities have been implemented by explicitly adding certain tags, by manipulating the “goto next” tag and adding calls to respective CGI executable to handle different types of requests (place anchor, recall anchor, etc.). These CGIs are really code transformers that take a VoiceXML document as input, and produce a VoiceXML document as output, with the desired feature added. Thus, as user navigate

dynamically, the VoiceXML document is constantly being regenerated, with the same textual contents but with different tags, to achieve the desired control feature.

When the VoiceXML documents are enhanced, two global variables, namely, 'currentName' and 'nextName' are added, that respectively store the id of the current form being played and the id of the next form to be played. By manipulating the values of these variable and the "goto next" tag during the regeneration process, the forms can be played in any sequence. This allows the feature of skip (skipping through paragraphs), repeat (repeating the current paragraph), start (reading the page from the start), and end (going to the end of the document). Another variable called "url" stores the URL of the current Web page in the VoiceXML document. When the user wants to go to any link on the current Web page, the URL of the current web page is stored on the stack (on the server, where the CGI is executed), hence when users says the keyword "back", the last URL on the stack is popped out and the page is retrieved. The forward keyword (the opposite of back) works in a similar manner.

Anchors or Bookmarks: In visual Web, users bookmark the pages, which they think has pertinent material. Those pages also have less important text which a user cares less about. In order to give users more control over the flow of the document and avoid impertinent text, we have introduced the concept of *voice anchors*. Anchors give the user the capability to select only those paragraphs which they want to hear, hence allowing them to have full control and have complex interactions, which they did not have earlier.

Currently Voice XML technology cannot interpret arbitrary utterance of a word from the user and convert it into text. Current Voice recognitions techniques are not powerful enough to recognize any arbitrary word spoken by an arbitrary user. It can only recognize words that are encoded in the *grammar* contained in the document in advance. In order to overcome this limitation, voice browsers support dynamic grammars such as GSL (Grammar Specification Language) or SRGS (Speech Recognition Grammar Specification). By dynamically generating the VoiceXML enhanced pages, we are able to incorporate dynamic grammar into the web page hence enabling dynamic voice recognition. Our system uses GSL grammar. GSL is a dynamic grammar, which grows over time. In GSL grammar, the user has to spell out the word character by character, but this only happens once. The next time the user needs to utter the same word, he/she does not have to spell the word again, as it is already added into the grammar. The grammar recognizes the words that it encodes. There are 26 letters and 10 digits recognized by our grammar. When a user spells a word, its characters are concatenated into a string and once the user completes the word, it is added into the grammar. After user's confirmation, it is stored permanently in the grammar. Our grammar is an offline grammar, which can grow indefinitely, can recognize words that exist in it, and can store any word possible.

Whenever a user wants to mark a paragraph, it places anchors on it. Our CGI is invoked with the values of currentName, nextName and anchor name. As mentioned earlier currentName and nextName are the global values, which gives

the unique current form id number and the next form id number of the document and anchor name is the user defined name. These values are stored in the offline memory through the CGI. Once the anchors are placed in the document, any kind of sophisticated navigation is possible. When the user wants to recall those paragraphs, another CGI is called, which regenerates the VoiceXML page with the control shifted to the paragraph that was marked by the anchor. If multiple paragraphs are marked with the same anchor name, they will be read out one by one in sequence (this feature thus allows subdocuments to be created on the fly). With Voice anchors, users can freely navigate through this enhanced document containing only the pertinent information.

5 Future Work

A prototype of the VoxBoox system is operational (for a demo, please visit the following URL http://www.utdallas.edu/∼gupta/demo.html). However, considerable work needs to be done. This work primarily involves adding features to the system so as to enrich the aural browsing experience of the listener. These features include:

- Obtaining summary of a paragraph or of the book: if a user wishes to obtain the summary of a paragraph, then they just utter the word "summary", at which point the first and the last sentences of the paragraph will be read out. Uttering "book summary" will yield the summary of all the paragraphs in the whole book.
- Storing subdocuments: As explained earlier, a user can mark multiple paragraphs with the same anchor name. When this anchor name is uttered, all marked paragraphs are read in sequence. Users can thus create hierarchies of subdocuments using this feature. At present, a subdocument created in this way cannot be stored and played again in a future session. We would like to give the user the ability to store the information about such subdocuments and retrieve them later. Such stored subdocuments can be conceivably made available to other users as well (e.g., instructor's collection of important passages of the document that the students should have access to).
- Creating an index on the fly: Most of the books come with an index, but many HTML-published books (especially novels and works of fiction) may not. For rapid navigation, we would like to allow the user to create an index by just giving a simple voice command. The index will be a VoiceXML page that will be synthesized from the VoiceXML book and played to the user. The user can then use the index to navigate through the book.
- Implementing voice-commanded scripting language [10] on top of our system to allow users to program simple navigation strategies using voice. For example, a user can mark a passage using the voice anchor "plot," another passage using voice anchor "climax", and then verbally program the navigation strategy "repeat climax after plot until stop" which will cause the passage marked "climax" to be played after the passage marked "plot" in an infinite loop, until the user utters "stop."

6 Conclusions

In this paper we presented the VoxBoox system: a system for making books published using HTML on the Web and making it accessible to visually impaired individuals via voice and audio. The advantage of our system is that it leverages existing standardized technology and Web infrastructure. It also allows dynamic, user-controlled aural navigation of the book. The advantages of our system over existing approaches (audio cassette books and digital talking books by DAISY) are the following: (i) navigation is via voice commands, rather than manual (pressing keys of the cassette player or using computer keyboard); thus, to use our system, one does not need a computer: a cell phone or a land-line phone suffices. (ii) New books can be rapidly added to the collection of accessible books, since all our system needs is for the book to be digitally available in HTML on the Web. Thus, as long as the book is published or composed digitally, and the publisher can provide an HTML markup of the book, the book can be made immediately accessible. (iii) More sophisticated aural browsing is possible, compared to audio cassettes and DAISY's digital talking books.

Acknowledgments. Authors have been partially supported by NSF, Dept. of Education and EPA. We thank Dr. Dipendra Manocha of National Association for the Blind, New Delhi, India, for suggesting us to build the VoxBoox system.

References

1. Bingham, H.: Digital Talking Books Expanded Document Type Definitions (2002) http://www.loc.gov/nls/z3986/v100/dtbook110doc.htm
2. Cookson, J. et al.: Digital Talking Books: Planning for the Future (1998) http://www.loc.gov/nls/dtb.html#activity
3. The DAISY Consortium http://www.daisy.org/about_us/default.asp
4. American National Standards Institute. Specification of the Digital Talking Book (2002) http://www.niso.org/standards/resources/Z39-86-2002.html#Strategy
5. Nichols, M., Wang, Q., Gupta, G.: A VoiceXML-based Spoken Scripting Language for Voice-based Web Navigation. In: Human Computer Interaction Conference, July 2005, Lawrence Erlbaum, Mahwah (2005)
6. McGlashan, S., et al. (eds.) Voice Extensible Mark Language (Version 2.0) http://www.w3.org/TR/VoiceXML20/
7. Gupta, G., Sunderraman, S., Nichols, M.: DAWN: Dynamic Aural Web Navigation. In: Proceedings of, International Conference on Human Computer Interaction (2005)
8. Annamalai, N., Gupta, G., Prabhakaran, B.: An Extensible Translator for translating HTML to VoiceXML. In: Proc 9th International Conference on Computers Helping People. LNCS, vol. 3118, pp. 339–346. Springer, Heidelberg (2004)
9. Reddy, H., Annamalai, N., Gupta, G.: Listener-controlled Dynamic Navigation of VoiceXML documents. In: Proc 9th International Conference on Computers Helping People. LNCS, vol. 3118, pp. 337–354. Springer, Heidelberg (2004)
10. VoiceXML Review http://www.voicexmlreview.org

An On-Screen Keyboard for Users with Poor Pointer Control

Rick Kjeldsen

IBM T.J. Watson Research Center, P.O. Box 704, Yorktown Heights, NY 10598
fcmk@us.ibm.com

Abstract. This paper describes a novel design for an on-screen keyboard that makes it easier for some users with relatively poor pointer control to type. The keyboard uses the edge of the screen to reduce the dimensionality of the pointing task required to identify a character, and uses gesture, rather than click or dwell to select it for typing. A prototype keyboard is used in a study involving people with and without physical impairments using several types of pointing devices. The results show that this approach has some significant advantages compared to a conventional two dimensional on-screen keyboard, especially with some types of pointing devices.

1 Introduction

Users who have Cerebral Palsy, Multiple Sclerosis, spinal cord injuries or similar motor impairments often have difficulty using physical computer input devices like a keyboard or mouse. For many, moving the pointer is possible, using devices including head trackers, joysticks, trackballs, touchpads or even standard mice, but typing on a physical keyboard is very difficult. These users often rely on on-screen keyboards for textual input.

While the design of on-screen keyboards for mobile devices has benefited from a great deal of innovation recently, the design of keyboards for users with physical impairments seems to be dominated by older designs that arrange characters in a two dimensional grid patterned after their physical counterparts. In our experience, this approach is not ideal for many users.

Standard on-screen keyboards assume that the user has sufficient accuracy of pointer control to easily identify a small two dimensional region of the screen, and the ability to click while hovering at one spot. Users with physical impairments often can control a pointer only coarsely. It may be difficult for them to keep the pointer still, and fine motor coordination may be insufficient to move it accurately over small distances. Once the pointer is on a character, a user may have trouble clicking to type it without moving the pointer. Often users will click inadvertently as they move about the screen, causing unexpected and undesirable actions. As an alternative to point&click typing, point&dwell is often provided. Here the user must keep the pointer still for a dwell period, after which a click is automatically generated. If a user has trouble keeping the pointer still, as might be the case with a head tracker for

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 339–348, 2007.
© Springer-Verlag Berlin Heidelberg 2007

example, this pause can be a problem. Waiting for each character in this way also makes typing very slow.

We have developed an on-screen keyboard that attempts to address these issues. The characters are placed along the edges of the display and normally hidden. A user types a character by moving the pointer to the edge of the display to pop-up the keyboard, moving along the edge to select a character, then moving away from the edge to type that character.

By placing the characters along the edge of the display, we remove one dimension along which the user needs to accurately position the pointer. This can be thought of as "bracing against" the side of the display. Just as it is by bracing the heal of your hand against a table you can position a pencil to write far more accurately than by holding your hand floating above the paper, so we propose that by bracing against the edge of the display a user can position the pointer in one dimension easier and more accurately then by positioning in two dimensions simultaneously. Of course, this technique is based on the assumption that the pointer will only move to the edge of the display, and not wrap or move to another screen if the user moves the mouse further in the same direction.

Second, the user need not perform a separate action (e.g. pressing a mouse button or other switch) to type the character. By using the same modality for typing as for selection, we reduce the number of different inputs a user must provide, hopefully reducing the complexity of the interaction. At the same time, moving away from the screen edge to type is quicker than waiting for a dwell time to elapse, allowing users with better dexterity or more experience to type faster.

This keyboard was originally developed in conjunction with a head tracking pointer. In that context this style of typing is a natural fit. It is easy to “overdrive” the pointer against the edge of the screen so that it does not pull away (assuming the system does not readjust when pushed against the side of the screen, as some do), and with no inherent mechanism to generate a click, typing by pulling away from the edge is more natural than dwelling over a key. If the keyboard is on the top edge, the typing motion is to tip the head back to open the keyboard, aim left or right to select a character, then tip the head down again to type. In short order this sequence feels very natural and fluid. We have since begun to explore the suitability of this approach for other input devices, and this paper reflects some of that work. We will describe the keyboard, and the results of a study involving users with and without physical impairments using various pointing devices.

2 Related Work

A number of commercial on-screen keyboards are intended for users with physical impairments, for example .the Grid (www.zygo-usa.com/grid.html), Madentec’s Discover Screen (www.madentec.com) or WiViK (www.wivik.com) to name just a few. While these focus on the needs of users with physical impairments, most do it without straying far from the traditional key layout (a 2D grid in QWERTY or alphabetic order). These systems generally support various typing styles including point&click, point&dwell, and scanning, but the design of the keyboard itself does not take into account the issues faced by a user who has difficulty pointing accurately.

Work that does address the design of the keyboard itself is usually focused on configuring the layout of the keys for efficient typing, generally in consideration of Fitts' Law [8]. By altering the spacing and layout of the keys, the time to move between them can be minimized, improving typing speed [7]. Unfortunately there is some evidence that such an analysis may not apply well to users with physical impairments such as Cerebral Palsy [1].

Interesting work has been done on the design of on-screen keyboards for mobile devices. Most of these are designed for stylus input, but some are relevant in this context. In [3] Mankoff and Abowd present a keyboard that types as the pointer enters keys laid out in a circular pattern. This allows users to type using gestures similar to cursive writing, as they move from letter to letter, but it requires a high degree of precision and so is not well suited to users who have difficulty accurately manipulating the pointer.

Wobbrock has designed the EdgeWrite system which he has adapted for PDAs, and more recently Trackballs [6]. This system allows a user to "type" by sketching approximations of letters around the edges of a square, or in the case of the trackball implementation by "pulsing" the pointer in the appropriate directions to simulate the same interaction. The system is well suited to people with poor pointer control using physical input devices, because the interaction can occur within a physical rectangle which helps constrain the user's motions. It is not clear how well this approach would work with a non-contact pointer control method such as a head tracker.

Dasher[5] is an interesting alternative to a standard keyboard where users select letters by moving the pointer into expanding colored regions which travel across the screen. This makes typing a continuous gesture, however the visual presentation is always changing, which can make it intimidating to novices and cognitively impaired users. It is also not clear how the system will perform for a user with poor hand-eye (pointer-eye ?)coordination.

3 Keyboard Description

The ScreenEdge Keyboard is simple in concept. Whenever the system pointer moves to the top edge of the screen, the characters in the alphabet appears along the top edge (Figure 1). The user moves the pointer left and right to highlight the desired character, then down from the top of the screen to type it. Just as with a traditional keyboard, the active window receives the character.

In the prototype, this primary keyboard area holds the characters of the English alphabet (A-Z), plus BackSpace. By default these characters fill the top edge of the screen. The size of the keys can be enlarged horizontally to make them easier to select. Since only a portion of the keyboard is now visible, the keyboard scrolls when the pointer reaches the edge of the screen (Figure 2).

In order to provide more space for characters, when the pointer reaches the left or right edge of the keyboard (i.e. enters the "<" key on the left edge or the ">" key on the right edge – the user may have to scroll to get there) another keyboard opens along the edge of the screen (Figure 3). This keyboard is navigated by traveling up and down along the edge of the screen, and characters are typed by moving away

from the edge. These side keyboards can also scroll up and down, though in the current versions they need not. The keys on the bottom of these side menus are Space and Enter so that the user only need move the pointer to the top corner, then straight down for these common operations.

Fig. 1. Keyboard in use with a head tracking pointer

Fig. 2. Screen Edge Keyboard, enlarged for better readability, scrolls when the pointer reaches the edge

The current prototype has been optimized to support web browsing, especially the typing of URLs. Toward this end, the right menu contains several shortcuts, such as single keys for “ www.” and “.com”. There is currently no provision for capitalization, and not all the keys on a traditional keyboard are supported. These shortcomings will be addressed over time.

This basic algorithm worked very well with a head tracking pointer, for which it was originally designed, but when using the keyboard with a mouse or trackball, new users would often type characters inadvertently. With a head tracker or joystick, it is easy to pin the pointer against the top edge of the screen as it is moved back and forth

Fig. 3. When the pointer reaches the left or right of the keyboard, a menu appears down the side of the screen

to select a character. With a mouse or trackball, however, no matter how far up the pointer is pushed, a very small down movement will take it away from the screen edge again – typing a character. To complicate matters, with a mouse, moving the pointer straight left or right without moving it down at all is difficult. As a result, users have to continually move the pointer up as they moved left and right, to keep it against the top edge, which causes the mouse to move further and further from the user as they work. This is an inconvenience for unimpaired users, but it is a real problem for users with a limited range of hand movement. Early trials showed that because extending their arm is difficult, many such users move a mouse left/right by swinging their arm left and right about the pivot of their elbow. As the mouse moved in this arc, it would pull away from the top edge, typing a character.

To improve performance with mouse pointers, we modified the behavior of the keyboard so that it captures the pointer within it till a character is typed. Instead of simply looking for any motion away from the screen edge, the system now types only when a pointer movement is above 55 degrees relative to the edge. If the motion is less than this angle, the pointer is captured within the keyboard as it moves, meaning that the component of movement parallel to the screen edge is retained, but the component perpendicular to the edge is forced to remain within the key board. The net effect is that typing requires a more precise movement directly away from the screen edge, and until this occurs the user can only move left and right within the keyboard. To allow users to exit the keyboard without typing, if the pointer pauses for more than 3 seconds the keyboard will disappear. A movement directly away from the edge will then exit, and restore normal pointer movement. A movement along the edge will re-open the keyboard and re-capture the pointer.

These changes greatly decreased the number of falsely typed characters. Users with poor pointer control are now captured within the keyboard till they type or pause long enough to exit.

To provide visual feedback we use color. When a key is selected, as the pointer moves back and forth along the edge, it turns green. After a character is typed, it changes to blue. This helps confirm the typing to the user and avoids the need for them to glance down at the window to see the character. After typing, the keyboard remains visible for 2 seconds, both so the user can see the highlight, and to allow

them to begin targeting the next character if there will be one. Some users have requested that this time be increased so that the keyboard remains visible longer.

4 Experiment

To evaluate the performance of the ScreenEdge Keyboard (SEK) it was compared to a traditional on-screen keyboard during several days of testing. Because the keyboard has been developed on Linux, the comparison keyboard was the Gnome On-screen Keyboard (GOK) version 1.0.3. Users were asked to type 3 sentences on each keyboard and the average character typing time was used to compare the keyboards.

Twenty three users participated in 29 evaluation sessions. Of these, we were able to get sufficient data from 17 users to make valid within-user comparisons between the keyboards. Only data from these users are included in the following discussion.

Eleven of the 17 were clients at United Cerebral Palsy of Suffolk, NY. All had previous computer experience using a mouse or trackball as a pointing device. All had limited hand and arm movement and used the pointing device with various levels of difficulty. All were capable of using a standard (physical) keyboard, and none had prior experience with on-screen keyboards. Because we did not have time to train them in the use of a head tracker, they used the input device they were most used to. One chose to use a trackball, the remainder used a standard computer mouse.

Six participants had no physical impairment. These were either staff at UCP, peers at IBM Research, or various family members. All had extensive computer experience. All of these participants were asked to use a head tracker [2] as an input device, three had sessions with other devices as well. None had previous experience with head tracking pointers. Two had prior experience with on-screen keyboards.

Each user was given a demonstration of the two keyboards, and then given as much time as they liked to practice with them before timing data was recorded. When they felt ready, they were then asked to type three sentences with each keyboard (6 different sentences per user). The sentences were chosen from a pool of 15 sentences such as “My favorite color is blue” and “It is almost time to go home”. The keyboards were used in semi-random order, either taking all data first from one keyboard then the other, or alternating keyboards with each sentence.

Typing in SEK was as described earlier in this paper. SEK was configured to fill the top of the screen without scrolling. Typing in GOK involved clicking on the keys of the Composition window (Figure 4). It can be argued that using GOK with a point&dwell strategy may be a better comparison, as SEK does not require clicking. We decided, however, that the dwell time required to type a character (500-1000 ms per character) would unfairly bias the data toward SEK. Since our subjects were generally able to click well, we chose to use GOK with point&click typing. The keys in GOK were set to be the same width and height as the width of the keys of the SEK keyboard. The characters were arranged alphabetically.

Timing data was recorded by a custom program which recorded the time each keystroke event was received with millisecond accuracy. The users were able to see the characters as they typed them into this program’s window.

The results of the sessions are summarized in Table 1. Over all sessions, SEK was about 1 second faster in terms of time-per-character (TPC). The fastest average TPC for one user was 1.9 seconds for a user with no impairments using a trackball on SEK. The slowest average TPC was 23.4 seconds for a user with CP using a mouse on GOK, who had great difficulty keeping the mouse still while clicking. Excluding this user, the highest average TPC was 11.9 seconds for a user with CP using a mouse on GOK.

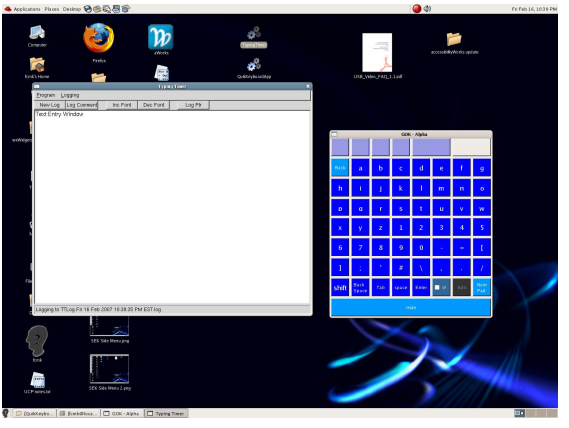


Fig. 4. The Gnome On-screen Keyboard (GOK) Composition window

Table 1. Results of SEK field trial. Times are in seconds / character typed. Note that the absolute times between lines can not be compared because of different sets of users on each line, and the large variability in typing ability between users. Relative times on each line are important.

Line	Group	SEK time	GOK time	Gain w/ SEK
1	Everyone	4.9	6.0	17%
2	User with CP	6.0	7.3	18%
3	Unimpaired users	3.4	4.1	16%
4	CP users except J	5.9	5.5	-7%
5	CP users – 1st session	5.7	4.8	-20%
6	CP users – 2nd session	6.7	8.3	19%
7	All Mouse users	5.4	5.2	-4%
8	All Head Tracker users	4.4	6.2	28%

To really understand the merits of SEK, it is informative to examine the data more closely. The first cut is to separate the data for users with and without motion impairments. Lines 2&3 of Table 1 show the TPC for users with CP versus non-impaired users. In both groups, SEK proved to be about 15% faster than GOK. Looking at the individual session data, however, many of the users with CP were slower with SEK

than GOK. This observation fits with comments we received from many of these users that they felt awkward using SEK.

The data for the users with CP was greatly influenced by one user, "J", mentioned above, who had a very large variation between his times with the two keyboards (SEK: 6.9 seconds, GOK: 23.4 seconds). The cause of this was that when he attempted to click on the GOK key, the mouse would often move inadvertently. This is a common issue with some types of disabilities [4], and highlights one of the advantages of the gesture-based typing action used in SEK. This large discrepancy, however, does hide the general trend in favor of GOK. In Line 4 we see that without J, the users with CP were about 7% slower with SEK.

Three of the users with CP participated a second time after using SEK occasionally for one week. To examine the effect of this practice, Line 5 shows users with CP on their first exposure to SEK, and Line 6 shows the average TCP for users on their second session. Here we clearly see that the first time users with CP are about 20% slower using SEK, but after practice the affect reverses. All the repeat users also reported liking SEK better than GOK in their second session. (Note that the longer absolute times on line 6 are because only a subset of the users had a second session)

Users with no physical impairments were universally faster with SEK, regardless of the device they used (mouse, trackball, touchpad or head tracker). They also reported preferring SEK in most cases.

Finally, we examined how various input devices performed with SEK. Lines 7&8 show the average for all mouse users, and the average for all head tracker users (note the mouse users include impaired and non-impaired users, the head tracker users were all non-impaired). The mouse users show about equal performance between the two keyboards. With the head tracker, however, SEK is significantly faster than GOK. This corresponds to the comments received from our participants. All 6 users who used the head tracker reported a preference for SEK. Interestingly, two users who used other devices with SEK for the first time reported preferring GOK, even though their typing times were faster with SEK.

5 Discussion

The numeric results from the study show clearly that the ScreenEdge Keyboard is both faster and preferred when used with a head tracking pointer. Results for other devices are not so clear cut. Users with no physical impairments, and experienced users universally typed faster with SEK using a mouse. First time users with CP were universally slower using a mouse with SEK, and reported liking it less. We suspect this is due to the difficulty these users had learning the gesture-based interaction. Many of them have worked very hard to achieve some skill with point&click interaction. To compensate for their poor pointer control, they have often developed idiosyncratic pointing styles. While a well coordinated user will generally move directly to a target in a ballistic movement, then home in with a short feedback driven phase, many users in our study take less direct paths to the desired target. One good example was S, who would uses a series of curving paths shaped almost like a fish hook, that would get progressively closer to the target. Once inside she would carefully click the

mouse button. This suggests that in spite of being able to point and click successfully, these users had not learned to accurately control the path of the pointer in the process. Because typing in SEK depends on the path of the pointer, it was initially more difficult for them to use. There is evidence that given sufficient time these users can perform well with SEK, but it remains to be seen how many users would have trouble learning to control the pointer path sufficiently to type well.

The experience of J (see above) suggests that the gesture-based interaction of SEK is a big advantage for users who have trouble clicking. Because he would move the pointer when clicking, his time per character with GOK's point&click interactions was almost 4 times longer than with SEK. We expect there would be a similar advantage for users who can not click at all, for if dwell times were added to the results, SEK would have bested GOK far more often. Some users, however, reported difficulty NOT clicking when using SEK, and in one case this clearly slowed her down.

Several users commented on the simplicity and lack of clutter with SEK. Having the keyboard appear on the edge of the screen when needed made it far less obtrusive than the large 2D keyboard window.

The most praise for SEK came from those who used it with a head tracker. They liked the nod-to-type motion used for typing characters on the top of the screen. One user reported she felt she was able to anticipate the pointer movement across the top of the screen, and begin the downward "typing" movement as the pointer moved into the desired character, making for a very fluid typing rhythm. Typing on GOK with a head tracker required a dwell, and that pause tended to break her train of thought.

6 Conclusion

We have described a new style of typing for on-screen keyboards where the characters are arranged on the edge of the screen, and the user types by moving to the edge of the screen, sliding along it to the desired character, and moving way again. For users with physical disabilities, this approach has some advantages. Users who have difficulty clicking can type without incurring the overhead of dwelling on each key. The character selection and typing motions form a gesture which users reported as fluid and natural. Users who have difficulty keeping the pointer stable can force it against the edge of the screen, and only need to precisely position the pointer in one dimension during character selection, potentially making character selection easier

Experimental data suggest that while the method works best with head tracking pointers, even with other devices, typing with the screen edge keyboard is faster than with 2D on-screen keyboards once the user has become comfortable with the movement pattern. Users with some types of motor impairments, however, may take some time to get comfortable with it.

Acknowlegements. I would like to thank the staff and clients of United Cerebral Palsy in Suffolk County Long Island for their gracious hospitality and help in testing SEK, especially on such a tight schedule. I would also like to thank Shari Trewin for her thoughtful comments, suggestions, and hours typing - and watching others type - simple sentences.

References

1. Gump, A., LeGare, M., Hunt, D.L.: Application of Fitts' law to individuals with cerebral palsy. Perceptual and Motor Skills 94(3), 883–895 (2002)
2. Kjeldsen, R.: Improvements in Vision-based Pointer Control. In: Proceedings of ASSETS 2006: 8th International ACM SIGACCESS Conference on Computers and Accessibility, Portland, USA, ACM Press, New York (2006)
3. Mankoff, J., Abowd, G.D.: Cirrin: a word-level unistroke keyboard for pen input. In: Proceedings of the 11th Annual ACM Symposium on User interface Software and Technology UIST '98, pp. 213–214. ACM Press, New York (1998)
4. Trewin, S., Keates, S., Moffatt, K.: Developing Steady Clicks: A Method of Cursor Assistance for People with Motor Impairments. In: Proceedings of ASSETS 2006: 8th International ACM SIGACCESS Conference on Computers and Accessibility, Portland, USA, pp. 26–33. ACM Press, New York (2006)
5. Ward, D.J., Blackwell, A.F., MacKay, D.J.C.: Dasher –A data entry interface using continuous gestures and language models. In: Proc. UIST '00, pp. 129–137. ACM Press, New York (2000)
6. Wobbrock, J.O., Myers, B.A.: Trackball Text Entry for People with Motor Impairments. In: Proceedings CHI'2006: Human Factors in Computing Systems. Montreal, Canada pp. 479–488 (2006)
7. Zhai, S., Hunter, M., Smith, B.A.: Performance Optimization of Virtual Keyboards. Human-Computer Interaction 17(2&3), 229–269 (2002)
8. Zhai, S., Kong, J., Ren, X.: Speed-accuracy trade-off in Fitts' law tasks — On the equivalency of actual and nominal pointing precision. International Journal of Human-Computer Studies 61(6), 823–856 (2004)

Multimodal Technology for Municipal Inspections: An Evaluation Framework

Irina Kondratova[1], Jeff Rankin[2], and Ashley Goggin[2]

[1] NRC Institute for Information Technology, National Research Council Canada
46 Dineen Drive, Fredericton NB E3B 9W4, Canada
Irina.Kondratova@nrc-cnrc.gc.ca
[2] Department of Civil Engineering,
University of New Brunswick, Fredericton, Canada E3B 5A3
{rankin,i8r19}@unb.ca

Abstract. Our paper discusses research on using multimodal interaction to improve usability of mobile data entry. This can benefit users of mobile technologies such as mobile workers, construction crews, and students collecting data during field trips. The evaluation of a mobile multimodal application for construction field work was conducted in the laboratory. Designing this laboratory experiment required a detailed definition of the type of device to be used, user requirements, including interaction techniques and usage context. During this work we discovered a need to develop a generic evaluation framework to assist in the selection of mobile technology and interaction techniques for further testing. The development of such a framework for municipal inspections with a focus on field inspections is discussed. This paper presents examples of using our framework to gain a better understanding of the contextual aspects influencing the feasibility of using mobile technology in the field.

Keywords: mobile technology, multimodal interaction, speech recognition, usability.

1 Introduction

Recent surveys of mobile Internet users show that usability is by far the biggest source of frustration among the users [1]. In particular, for learning applications, research shows that the most important constraining factors for widespread mobile learning (m-learning) adoption, along with battery life, are the screen size and user interface of most portable devices [2]. These usability problems are exacerbated when mobile devices are being used in the field for data collection and information management, where, in addition to above mentioned usability problems, users are constrained by the environmental conditions, including lighting, noise, weather, etc. This directly relates to the feasibility of using mobile devices for data collection by students during field trips.

Field trips play a vital role in engineering and science education, allowing students to acquire valuable knowledge and experience in the field while using field

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 349–357, 2007.
© Springer-Verlag Berlin Heidelberg 2007

instruments and learning efficient data gathering techniques. Handheld devices have immense potential to be utilized during these field trips as data gathering devices, allowing the information collected in the field to be saved for future analysis in class [3]. Using handheld wireless devices in the field also enables real time communication with other groups of students as well as with a learning repository.

In order to assist users in managing mobile devices, user interface designers are starting to combine the traditional keyboard or pen input with "hands free" speech input [4], adding other modes of interaction such as speech-based interfaces that are capable of interpreting voice commands [5]. As a result, speech processing is becoming one of the key technologies for expanding the use of handheld devices by mobile users. In a recent eLearning technology foresight, technology-based education guru Tony Bates predicted that: "A new computer interface based on speech recognition will have a major impact on the design of e-learning courses" [6]. Currently, automated speech recognition (ASR) technology is being used in desktop e-learning applications for automated content-based video indexing for interactive e-learning [7], audio–clip retrieval based on student questions [8], and, together with speech synthesis, to improve accessibility of e-learning materials for visually impaired learners [9].

Speech technology seems to be ideally suited for enhancing usability of m-learning applications designed for the mobile phone, where speech is a natural way of interaction and where a small screen size limits the potential for a meaningful visual display of information [10]. However, speech technology is limited to only one form of input and output - human voice. In contrast to this, voice input combined with the traditional keyboard-based or pen-based input model permits multimodal interaction in which the user has more than one means of accessing data in his or her device [11].

Multimodal interfaces allow speedier and more efficient communication with mobile devices, and accommodate different input modalities based on user preferences or on the usage context. For example, in a field trip scenario a civil engineering student would request information about the field structure (bridge, building, road, etc.) from the course repository using "hands free" voice input on a "smart phone" (hybrid phone-enabled PDA). The requested information would then be delivered as a text, picture, CAD drawing, or video, if needed, directly to the PDA screen. The user will be able to enter field notes in the forms using a portable keyboard or a pen, if appropriate, or via voice input during field data gathering. In addition to this, verbal field notes could be attached to the data collected in the form of an audio recording.

To evaluate a particular type of mobile technology for field work and choose appropriate modes of user interaction with a mobile device, technology evaluations are conducted in the field and in the laboratory. Laboratory evaluations of mobile technology use permits tightly controlled usability studies while simulating real-life usage environments [12]. To re-create field usage environments in the laboratory, in-depth preliminary research should be conducted in order to single out a set of the most important parameters affecting technology use in the field. The success of laboratory testing often depends on the ability of a researcher to properly capture and recreate these contextual usage parameters.

2 Laboratory Evaluations for Mobile Field Data Entry

In our ongoing research work on applications of mobile devices in the field for data collection and field notes taking, we developed a mobile application that allows a concrete inspector on the construction site to enter field testing data. Using this application, a concrete inspector can document the results of field concrete quality control tests using various interaction modes such as speech, stylus and a soft keyboard on the handheld device [13]. The application was evaluated in the laboratory by a group of users.

The evaluation was designed to recreate field conditions for concrete inspectors on the construction site. The typical construction site noise was recreated by conducting the study under noise conditions within a range of a typical construction site. The audio used to recreate the noise was recorded during concrete placing at a real construction site, thus providing a realistic evaluation environment for a mobile application that is designed to support data entry by concrete technicians. The experimental set-up insured a certain amount of mobility requiring a participant to travel from one table to another, a procedure that is similar to a concrete inspector performing quality control testing at the location of one concrete pour and then moving to another location on site [14].

To measure the awareness of participants of the surrounding environment during data entry, researchers projected random images on the walls of the lab and participants were asked to count certain types of images. A "safe" typical construction image depicted a building construction site with no heavy machinery present. A "danger" image projected was an image of a large concrete truck that field workers should be aware of while on site. During the experiment, participants were asked to count "danger" images shown while entering data or moving between data entry stations. As a result of this evaluation, we drew certain conclusions on the feasibility of using multimodal (speech and/or stylus and soft keyboard) data entry options in these conditions [15].

Designing this experimental evaluation in the laboratory required a significant amount of preliminary research on the type of device to be used, user requirements, including interaction techniques, usage context, etc. Therefore, through our experimental design and technology evaluation process we discovered a need for the development of the generic evaluation framework that will allow us to rapidly evaluate and select mobile technology and interaction techniques to be used for mobile data entry so that we can conduct follow-up laboratory and field testing.

3 Mobile Municipal Technology Evaluation Framework

The authors initiated the development of such a framework in a research project with the City of Fredericton (Canada) as a research client. The city's Information Technology Department wanted to conduct evaluations of several types of field municipal inspections in order to choose a mobile solution for a broad variety of inspections.

The project's primary deliverable was the development of a method for selecting various technology features for a given municipal inspection type. Six inspections

from several of the municipality's departments served as the case studies in the development of the framework. The choice of inspections allowed representation of a broad range of possible contextual aspects and user requirements. Using this set of inspections as a case study, the aspects of an inspection that influence the selection of different forms of technology were determined, along with several conclusions on how mobile applications of technology may potentially improve an inspection process. Recommendations were developed for each individual inspection as well as an overall recommendation for a base mobile technology to be adopted by the municipality that would best fulfill the requirements of the majority of inspections.

3.1 Methodology

The mobile technology evaluation framework developed within this project can be used to evaluate mobile technology for a multitude of field applications. A block diagram of our framework is presented in Figure 1.

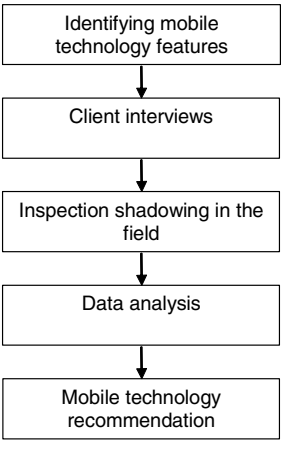


Fig. 1. Steps in the technology evaluation framework

Within the project we identified a number of technology attributes that should be taken into account when evaluating any mobile technology for field data entry applications. These attributes address specific requirements that clients may have as well as those factors derived from the field data gathering process, such as the information type and the surrounding environment.

The overall approach we have taken consisted of identifying certain perspectives of each individual inspection, data collection and analysis based on the templates we developed and recommendations on the technology solutions. Three perspectives of each individual inspection were defined as:

1. process steps for each individual inspection
2. information requirements
3. contextual aspects

Process steps provided a breakdown of several distinct steps common among the majority of inspections to provide a comparative basis during analysis. The second perspective of information requirements detailed the type of information captured or accessed, which may range from textual to high resolution images. Several key issues affecting the selection of technology were identified within the third perspective, contextual aspects. For instance, environmental factors such as the weather would influence the feasibility of certain equipment types and may require a waterproofing feature, or excessive noise on the site might not permit a reliable speech data entry option for inspection data.

3.2 Client Interviews

Information for analysis was collected through a series of interviews with the client's primary contact as well as with the department supervisor for each of the respective inspections. Each participant was asked a set of pre-defined questions to gather such information as: their personal level of knowledge and experience with mobile and multimodal technology: which inspection they would like to have studied: identification of specific contextual aspects of the chosen inspection such as its location, duration, and any potential hazards that may be present: training or certification that may be needed for the inspector: and any possible improvements that they would like to make to the inspection process.

From the interviews it became evident that each participant had their own set of unique requirements and concerns to be addressed. Further, valuable data was gathered through a process of inspection shadowing to verify previously collected data and to gain further insight into all perspectives of a given inspection.

3.3 Inspection Shadowing

Although interviews with the department supervisors provided valuable information regarding the purpose of the inspection, further investigation was needed to discern the actual inspection process that takes place in the field. To accomplish this through an observational study, each inspector was shadowed during an actual inspection and was then asked a series of questions to determine their opinions of the inspection process and how various forms of mobile, and possibly multimodal, technology may be beneficial.

The questions we asked were designed to isolate specific environmental and contextual features of the inspection process that may impact the applicability of certain technology. The questions also assisted in determining what sort of tasks the inspector may be doing during the inspection. In addition to providing valuable information regarding the specific stages of several inspections, the shadowing process also provided a means of validating the participants' responses during the initial interviews. It was shown that an individual's responses regarding their own requirements and concerns may depend on their own specific responsibilities. Thus, by cross referencing data collected during the inspection shadowing process with

information collected during client interviews, we were able to better define user requirements and usage context for each individual inspection.

As an outcome of this project, we developed a comprehensive procedure for analyzing technology requirements for a given city inspection, gathered detailed information in the field, and developed a systematic framework for analyzing the requirements of a given inspection [16].

4 Application of the Framework

To demonstrate an application of this analysis framework, the following section will illustrate the process we used to choose the appropriate mobile technology for a municipal concrete inspector and, as an example, for a group of civil engineering students on a field trip to the curb inspection site.

The following information on the concrete curb inspection was collected by the researcher via client interviews and inspection shadowing [16]. Concrete curb inspection of a concrete pour for curb and gutter is conducted to maintain standards in the product and the contractor's performance. It could be carried out under various weather conditions, including light rain. The potential risks at the site include traffic, excessive noise from equipment or vehicles, high speed traffic, large and heavy equipment and possible chemical exposure from curing compounds.

The inspector is moving around the worksite, using their hands which may get wet or dirty, using various pieces of equipment, speaking with others, making observations, and taking measurements. Currently all inspection information is collected using a standard paper form. No data privacy concerns were expressed by the participants. The current preferred method of communication is using a cell phone. Significant findings during inspection need to be reported immediately to a supervisor. Participants felt that usability and functionality will be the most important attributes of the device. A desire was expressed to have a permanent database of inspection results to aid in the review process for payment purposes and to examine the quality of the concrete at a later date. Participants felt that the device needs to be relatively simple to use and that the current inspection form should be reflected by the device.

Based on this information, the following mobile solutions for the concrete curb inspections were recommended to the city (a summary of this recommendation is presented in Table 1). We recommended to the city that the first choice for the concrete curb inspection would be a device similar to a Pocket PC (or a Smart Phone to provide a voice communication option) due to its small size and its ability to enter information efficiently with a stylus. The device needs to have a glare proof screen and waterproofing and should have a wireless connection to transmit significant findings to the office in real-time. It is likely that the interaction with the device will be limited to the use of stylus for the majority of the inspection due to the excessive noise at the worksite. Speech input may be feasible when certain equipment is not in use and there is minimal traffic activity.

Table 1. Proposed mobile solution for a municipal concrete curb inspector

Technology recommendation		Features required	Advantages and disadvantages
Preferred	Secondary		
Pocket PC	Tablet PC	- Mobility - Ease of storage - Real-time connection for automatic submission of forms - Stylus for input - Potential for speech entry when machinery is not in operation - Rechargeable - Glare proof coating - Waterproof casing - Potential for recording of comments	An advantage of Pocket PC is that the inspector may remain mobile throughout the inspection. The device is small enough to be easily stored, which allows the inspector to use their hands for other tasks.

The advantages of the suggested technology are:

- Due to its small size, the device will allow the inspector to freely use hands and other equipment by being easily stored away in a pocket.
- The inspector will be able to move around the worksite without hindrance.
- The majority of the information collected on the current inspection form is checkbox, based on the selection of two or three pre-defined codes. Therefore, a device such as a Pocket PC with a stylus would be suitable for the inspection as the inspector would be able to select the appropriate code to convey the concrete curbing process' condition at each stage.
- A real time connection would allow for immediate submission of the completed inspection forms to the main office for review.
- A glare proof coating would allow the inspector to view the information in the field regardless of the lighting.
- Speech input may be advantageous as it would potentially allow the inspector to remain cognitive of the surroundings. However, speech would only be feasible as a data entry method when the heavy equipment is not in use and there is minimal traffic activity during certain sections of the inspection.

The main disadvantage in using a Pocket PC is its inability to record detailed written comments. Although the current inspection process does not require detailed comments to be recorded, comments would be an asset for the reviewers of the information to gain a better understanding of the condition of the concrete. However, a Pocket PC has an option of recording audio files with comments that could be attached to the inspection form.

The main advantage of the Tablet PC would be its ability to record written comments; a larger comments section may allow for a more complete description of the condition of the concrete. It is also important to note that the inspector must remain aware of the surroundings at all times due to the potential risks that are present. Therefore, a Tablet PC may also be beneficial in terms of a better awareness of the surroundings as the screen would allow displaying any images or information

clearly for an inspector to see, and allow them to pay more attention to the surroundings at the same time.

Contrary to the Pocket PC usage recommended for a municipal inspector, a preferred mobile solution for a group of civil engineering students on a field trip to learn some basics of concrete testing would be a rugged Tablet PC with a glare proof coating. A Tablet PC will allow collecting data and additional notes in the field for a course report, while at the same time allowing more attention to be paid to the surrounding environment to provide for student safety on the construction site. Based on the assessment of the environmental conditions, the interaction technique with the device has to be limited to the use of a stylus due to the excessive noise at the worksite particularly when the concrete is being poured (noise level up to 95dB). However, speech input may be feasible when certain equipment is not in use and there is minimal traffic activity. Speech input may also be advantageous as it will allow the users to remain cognitive of their surroundings, as well as take additional field notes. The majority of the information collected on the current inspection form is a checkbox, based on the selection of two or three pre-defined codes, thus a stylus could be used efficiently to enter this information. Wireless connectivity will not be a requirement for this field work since all field notes could be downloaded to a desktop computer in class to produce a field report.

5 Conclusions

In many cases mobile technology for business, learning and field applications is underutilized due to the failure to choose an appropriate technology for particular usage conditions. The technology evaluation framework we developed in the research project with the local municipality as a research client, could serve as a generic framework for choosing mobile technology for field data collection applications, including m-learning applications.

The project's primary deliverable was the development of a method for selecting various features of technology for a given municipal inspection. Analysis was performed based on the information collected from the various interviews with the participants and from the process of shadowing the inspections. Both methods served as valuable tools for acquiring the necessary information. Although the two approaches complemented each other and served as a means of validating the information, the vast majority of the data was collected from the shadowing process in the field. While in the field, a broad range of contextual aspects for each inspection were identified through observing the inspector performing the actual tasks. This provided a better understanding of how a mobile device may potentially be used in the field. Information collected on each individual city inspection process will be utilized further in the development of the laboratory evaluations of mobile solutions for field work.

Acknowledgments

The researchers acknowledge the contributions provided by the employees at the City of Fredericton. This research project was supported by the Natural Science and Engineering Research Council of Canada and by National Research Council of Canada.

References

1. Sadeh, N.: M-Commerce: Technology, Services, and Business Model. John Wiley & Sons, Inc. New York (2002)
2. Trondsen, E.: Global Perspective on Mlearning. Macromedia Higher Education Leadership Forum San Francisco (2005)
3. Kravcik, M., Kaibel, A., Specht, M., Terrenghi, M.: Mobile Collector for Field Trips. Educational Technology & Society 7(2), 25–33 (2004)
4. Wilson, L.: Look Ma Bell, No Hands! - Voicexml, X+V, and the Mobile Device. XML Journal (2004)
5. Sawhney, N., Schmandt, C.: Nomadic Radio: Speech and Audio Interaction for Contextual Messaging in Nomadic Environments. ACM Transactions on Computer-Human Interaction 7(3), 353–383 (2000)
6. Neal, L.: Predictions for 2002: E-Learning Visionaries Share Their Thoughts. eLearn Magazine. 2002 (1) (2002) 2
7. Zhang, D., Nunamaker, J.F.: A Natural Language Approach to Content-Based Video Indexing and Retrieval for Interactive E-Learning. IEEE transactions on multimedia. 6(3), 450–458 (2004)
8. Guenaga, M.L., Burger, D., Oliver, J.: Accessibility for E-Learning Environments. In: Miesenberger, K., Klaus, J., Zagler, W., Burger, D. (eds.) ICCHP 2004. LNCS, vol. 3118, pp. 157–163. Springer, Heidelberg (2004)
9. Jahankhani, H., Lynch, J.A., Stephenson, J.: The Current Legislation Covering E-Learning Provisions for the Visually Impaired. In: Shafazand, H., Tjoa, A.M. (eds.) EurAsia-ICT 2002. LNCS, vol. 2510, pp. 552–559. Springer, Heidelberg (2002)
10. De Freitas, S., Levene, M.: Evaluating the Development of Wearable Devices, Personal Data Assistants and the Use of Other Mobile Devices in Further and Higher Education Institutions. JISC Technology and Standards Watch Report: Wearable Technology (2002)
11. Jokinen, K., Raike, A.: Multimodality - Technology, Visions and Demands for the Future. In: Proceedings of the 1st Nordic Symposium on Multimodal Interfaces Copenhagen (2000)
12. Kjeldskov, J., Graham, C.: A Review of Mobile HCI Research Methods. In: Proceedings of the 5th International Symposium on Mobile Human-Computer Interaction (MobileHCI03) Udine, Italy (2003)
13. Kondratova, I.: Speech-Enabled Handheld Computing for Fieldwork. In: Proceedings of the International Conference on Computing in Civil Engineering ICCC 2005 Cancun, Mexico (2005)
14. Kondratova, I., Lumsden, J., Langton, N.: Multimodal Field Data Entry: Performance and Usability Issues. In: Proceedings of the Joint International Conference on Computing and Decision Making in Civil and Building Engineering Montréal, Québec, Canada (2006)
15. Lumsden, J., Kondratova, I., Langton, N.: Bringing a Construction Site into the Lab: A Context-Relevant Lab-Based Evaluation of a Multimodal Mobile Application. In: Proceedings of the 1st International Workshop on Multimodal and Pervasive Services (MAPS 2006) Lyon, France (2006)
16. Goggin, A.: Mobile Multimodal Technology for Inspections. University of New Brunswick. Fredericton (2006)

The Effects of Spatially Enriched Browsing Shortcuts on Web Browsing of Blind Users

Christos Kouroupetroglou[1], Michail Salampasis[2], and Athanasios Manitsaris[1]

[1] Department of Applied Informatics, University of Macedonia, Egnatia 156, 54006, Thessaloniki, Greece
kourou@teithe.gr, manits@uom.gr
[2] Department of Informatics, Alexander T.E.I. of Thessaloniki, 57400, Thessaloniki, Greece
cs1msa@it.teithe.gr

Abstract. The Web has turned out to be an interactive and complex information seeking environment where apart from text, layout and visual cues play a significant role in its efficient use. In this paper we present a semantic web based framework for communicating this visual meta-information to blind users who miss it. Based on annotations of web pages blind users are given a set of browsing shortcuts within a web page and this way they can move easily to the desired point of the page. In addition to this mechanism we enhanced these annotations with spatial information and we provide blind users with a spatially enriched variation of the browsing shortcuts mechanism. To test its effects on web browsing of blind users we conducted an experiment with 15 users participating. The results indicate that there is no significant improvement when using the spatially enriched variation of the shortcut feature but there is a decrease in time and keystrokes when the browsing shortcuts mechanism is used alone.

Keywords: Web accessibility, Semantic Web, browsing shortcuts, spatial information.

1 Introduction

At the early days of WWW development, accessibility for Visually Impaired (VI) people was mostly a text-to-speech problem since the web kept a rather textual form. However, the dominance of the desktop metaphor in web page design, as well as the use of the web as a platform to deliver different types of complex applications, transformed the web into a more interactive, highly demanding and complex information seeking environment. This development affected seriously blind users' web accessibility because apart from the text itself, the layout of a web page and the other orientation and navigation aids started to play a significant role in using web applications efficiently. In fact, nowadays the visual meta-information which is communicated to sighted users and considered crucially important for efficient information seeking, is mostly missed by blind users [3].

The main source of the problem is the lack of semantics in HTML. HTML does not support the advanced structures that are presented through visual representations as

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 358–367, 2007.
© Springer-Verlag Berlin Heidelberg 2007

first class objects. For instance, the concept of a site map aiming to alleviate user disorientation is not supported inherently in the Web, but it is only perceived by sighted users due to its visual representation. In an effort to enrich web pages we developed a semantic web based application framework [6]. The framework depends on the commitment of people participating in it and forming communities with specific goals. The three groups of people within the framework are:

- *Ontology creators*, who develop ontologies that will provide annotators with the necessary vocabulary in order to annotate web pages.
- *Annotators*, who annotate web pages according to users' needs.
- *End users* who benefit from the annotations using framework-aware agents.

In this paper the agent which is discussed is a specialized voice browser (called SeEBrowser) which utilises the annotations implementing a browsing shortcuts mechanism. However, in this paper we focus more on a spatially enriched variation of browsing shortcuts (SEBS) in comparison with the simple serialized access (SSA) and the simple browsing shortcuts (SBS) mechanism. Sighted users use spatial information of a web page's design in order to remember and use a web site more efficiently in the future. In this paper we wanted to investigate the effects of this spatial information on web browsing of blind users. In order to investigate these effects we conducted a within subjects experiment using 15 blind users. The object of the experiment was threefold: 1) To investigate whether information given by annotations about the existence of elements in web pages is already known or perceived through SSA, 2) To find out whether blind users can exploit and use efficiently spatial information given to them in SEBS and 3) to examine the effects of the 2 variations of the browsing shortcut mechanism in the efficiency of information seeking tasks.

2 Related Work and Rationale

A solution to problems caused by lack of semantics in HTML, poor web site construction and the inefficient serialisation mechanism of screen readers was proposed by Tagaki and Asawaka [10]. They based their solution on annotations of web pages that helped in their reconstruction through a transcoding server. A similar approach was suggested also by Huang & Sundaresan [5]. However, Hanson & Richards [4] argue that the specific transcoding server approach has certain drawbacks.

To overcome issues pointed out by Hanson and Richards we used some of the ideas that the Annotea project [1] is based on. The simplicity of the project and the ideas of *client-server architecture,* having the *annotations stored on a server* and using them through specific agents are key points to our approach too.

Thus, unlike the efforts using a transcoding server, our Semantic Web (SW) framework does not require any centralised mechanism that transforms the web pages, but the utilisation of annotations takes place client-side. This makes our framework a flexible, open, extensible and scalable annotation scheme, while at the same time it aims to maintain the non-centralised nature of the WWW. Marshall & Shipmann [7] present and categorise the various views and expectations raised by the

SW. The view closest to our approach is that of people using the SW as a way for metadata syndication, enabling the communication between various information sources and agents.

3 The Application Framework

The application framework is described in detail in [6]. Moreover the purpose of this paper is to describe the effects of a specific implementation of the framework and not the framework itself. Thus, in the following sections we will present in brief the three basic parts of it.

3.1 Ontologies

Ontologies are the first key point of our framework and they are produced by ontology creators who need to have some experience in knowledge representation and engineering. Ontology creators use a specially adapted ontology editor called ONAR [11] but any OWL compliant ontology editor can be used.

The ontology developed up to now for the purposes of our project, describes navigational and visual aids existing in web pages and not being able to be captured in the application level such as the ones described in [13]. These elements are common "landmarks" in many web pages and in web portals and search engines in particular such as menus, advertisements, results list, internet directory etc. Apart from classes, ontologies can also describe relationships between them and a very important one described in our ontology is the "Contains" relationship. It indicates a composite element in a web page, for example a results list that contains result items etc. Finally the spatial information is kept in a common attribute for all classes which is called "Location".

3.2 Annotations

In the second part of the framework, annotators use ontologies in order to describe elements in web pages. Annotations are stored in RDF/XML files which are uploaded to a server and kept in a RDF database through a web interface. In this process there is also an authentication stage for two reasons. First we could not prevent multiple annotations from various annotators since we wanted to keep the framework open. The decision of which annotations to use is up to the end-users. Second, it ensures that the feedback given from the users group to annotators can be directed to the right persons. Annotators use a GUI based annotation tool to produce annotations. They can easily select an element from a web page and assign it to a specific class from the vocabulary. Another characteristic of the tool is a mechanism that calculates the position and the area that the element covers in the web page and produces the spatial information kept for the location attribute. This information is coded in a series of numbers from 1 to 9 separated by commas. Each of the numbers indicates an area on the web page, according to the numeric keypad diagram, which the element covers (see Figure 1).

3.3 Utilisation

The last part of the framework has to do with the utilisation of annotations using the SeEBrowser which uses the Demosthenes speech composer [12]. SeEBrowser as any other voice web browser includes basic browsing functions such as going back and for in history pages, browsing and speaking modes e.t.c.

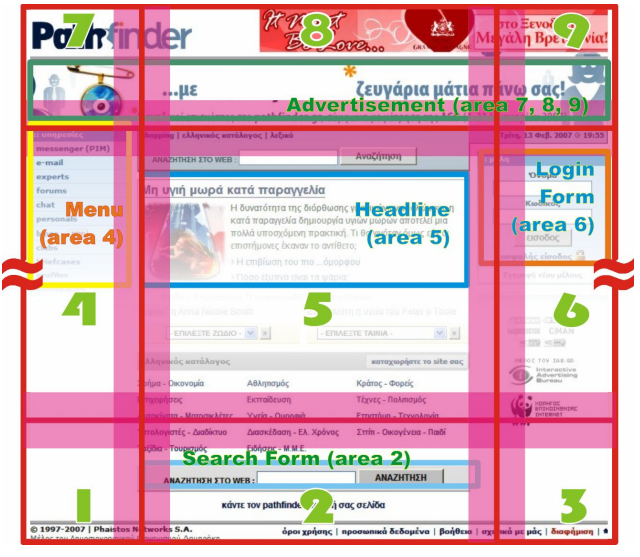


Fig. 1. Area separation and annotated elements allocation

However, the distinctive feature of SeEBrowser is the browsing shortcuts mechanism (SBS) and its spatially enriched variation (called SEBS). The SBS feature is presented in detail in [9]. Browsing shortcuts is a mechanism which facilitates blind people to move efficiently to various elements of a web page (e.g. functional elements such as forms, navigational aids etc.), hence operating effectively as an interaction method and a vital counterbalance to low accessibility of web pages. In general it provides blind users with a list of entry reading points to the web page that sighted users spot within seconds. Moreover, the feature enables rapid transition amongst parts of the page reducing the time needed to reach a specific point using the simple serialized reading of the web page. Finally, it also enhances a progressive disclosure mechanism that excludes annotated elements from the initial list when they are contained within others and presents them when the user reaches the container element.

The SEBS is a variation of SBS that, in addition to browsing shortcuts, utilises the spatial information kept in the annotations. In particular the user can listen to the annotated elements that exist in a specific area by pressing Ctrl+[Number] where [Number] is the number of the area according to the numeric keypad. In addition, there is also the option of listening to which areas an element covers by pressing Ctrl+0. This variation aims in making more efficient the use of browsing shortcuts

(SBS), based on knowledge of web design principles such as the desktop metaphor. For example, based on the desktop metaphor a menu of web site will probably exist in some peripheral area such as 4 of 8. Thus, when a blind user is looking for the menu on a web site s/he should first look in either of these two areas by pressing Ctrl+4 or Ctrl+8. This way s/he will have to browse through a much more limited list of annotated elements to find the desired one. However, blind users are not familiar in using spatial information since the serialized version of web pages does not include any. This means that they will probably need some time to familiarize, absorb and use it efficiently.

4 User-Centered Evaluation Experiment

In an effort to evaluate and investigate the effects of the spatially enriched browsing shortcuts we conducted an experiment following the experiments described in [6] [9]. These experiments gave useful and encouraging feedback regarding the efficiency of SBS and SeEBrowser's usability in general. However, in the experiment we report in this paper we investigate the SEBS feature in comparison to both the SBS and the SSA mechanisms. Thus, we made the necessary changes in the design of previous experiments to accommodate the new needs. Moreover, we wanted to determine whether spatial information given by the browsing shortcuts is already perceived without it and whether the spatial information is really understood by users.

4.1 Subjects

Subjects of this experiment were 15 blind users already using the web with screen readers in combination with Internet Explorer. All of them were rated as medium to expert users from their IT teachers in the school for blind people and they are using the web for at least one year.

4.2 Environment and Tasks

The environment used for the experiment consisted of 2 popular web portal sites (www.pathfinder.gr & www.metropolisradio.gr), used for the description tasks and 2 large newspaper sites (www.tovima.gr) (www.enet.gr) used for the information seeking tasks. Users were asked to complete 2 description and 6 information seeking tasks in overall.

The two description tasks of the two web portal home pages were completed using in one case the SEBS system and in the other using the SSA system. Each of the 2 cases covered a different goal. When using the SSA users were asked to identify and order four kinds of web page elements according to the serialization elements found on a web page. The elements to be identified were: Menus, advertisements, headline articles and groups of latest news. This task investigated whether information given by annotations about existence of elements in web pages is already known or perceived through SSA. In the second case users were asked to describe a "spatial map" of the web page as they where exploring it using the SEBS mechanism. This task helped to verify that blind users can understand the spatial information given to them in SEBS. The 6 information tasks were separated in 3 sessions. Each session

consisted of 2 question answering tasks. One task required within-page browsing and the second across-pages browsing in order to be answered and users were informed about it before each task. Each session was to be completed using a different variation of the system (i.e. SSA, SEBS, SBS). In the first two sessions subjects used each of the two newspaper sites. The third session was conducted after a period of about a month to avoid carry over effects from the use of the previous two systems and they used randomly one of the two newspaper sites which published new content at that time.

4.3 Protocol

A within-subjects laboratory experiment was conducted to compare the systems. The experiment was separated in three sessions. The first two sessions were conducted in a classroom having 2 subjects working simultaneously on the same task but with a different system. Each of these two sessions completed using a different system on a different web site. The two systems-conditions used in this first experiment period were the SSA and the SEBS. The sessions consisted of one description task and 2 question answering tasks. The descriptions of web pages from users were recorded in mp3 files.

Before the first session users were given a demonstration of the two systems and they were asked to use them for about 15 minutes for familiarisation. They were given a 30 minutes period for each task with short breaks between them and there was also a 30 minute break between the two sessions. The third session was conducted after about a month from the first two and consisted of two question answering tasks in one of the newspaper sites selected randomly for each user.

4.4 Instruments and Measures

The instruments used for the experiment analysis were the descriptions' audio files and the log files that kept their interaction with the system. The log files were keeping every movement they made along with information about the listening cursor at that time.

Descriptions. The information from the descriptions was recorded to either list of elements when the SSA system was used, and to 9-cell table maps when the SEBS system was used. The lists and maps were further analysed in order to group common problems found and point out specific cases required deeper analysis. These observations indicated the main effects of SEBS in the description of web pages.

Log files. The log files were used as instrument for the information seeking tasks. Based on them we calculated the following measurements/evaluation criteria:

- Number of keystrokes per task.
- Average number of keystrokes per visit for every task.
- Time for task completion.
- Average time spent per visit for every task.
- Average time per movement for every task.
- Times of repeated segments of a web page (text re-entries).

4.5 Results

Descriptions. The analysis of the descriptions of web pages showed some very interesting results. First of all, the description made using SEBS where most of them quite accurate and complete. However, one very common problem in these descriptions was that users could not understand when an element extended to multiple web pages areas. In many cases users didn't use the area identification (Ctrl+0) mechanism at all, which informs them about the areas that an element covers.

Furthermore, the findings from the descriptions using the SSA system are even more interesting. First of all, most of them are incomplete. The most commonly missed elements were advertisements and menus which are the most dependent on their visual representation elements out of the four requested. Missing alternative texts in advertisements banners caused confusion and menus are more often comprehended simply as consecutive links.

In addition, there were cases where users could not understand that they were reading latest news sections. They just read the article titles without realizing that they were sections with latest news of a specific domain. The least misunderstood and omitted element was the headline articles probably because of the short abstracts usually following them.

Time for task completion. As we can see in Table 1 the SBS system reduced the mean task time to 4.83 minutes instead of 7 minutes for the SSA. On the other hand the SEBS mean task time was 10.67 minutes. In fact, the paired samples T-Test proved that the differences between the following pairs (SSA - SEBS) and (SBS–SEBS) are statistically significant.

Table 1. Average results of evaluation criteria in three experiment conditions

Evaluation criteria	*System*		
	SSA	*SBS*	*SEBS*
Task time	7.00min	4.83min	10.67min
Avg. time in visits per task	3.37min	2.57min	4.62min
Avg. movement time per task	0.311sec	1,433sec	0.604sec
Keystrokes per task	237.24	152.17	280.20
Avg. Keystrokes per visits and task	123.76	87.77	144.43
Text re-entries	68.34	36.45	134.70

Average time spent per visit for every task. Table 1 shows that the same correlations between systems exist for the average time spent per visit for every task. The SBS system users spent less time per page than the SSA system users who in turn spent less time than the users with the SEBS system. In fact, the difference between the best (SBS) and worse system (SEBS) is statistically significant too.

Average time per movement for every task. The results in this metric can partially give some insight into the previous two evaluation criteria. We observe that in the SEBS condition users spent almost twice the time per movement than users in the SSA condition. In fact the differences between the systems are statistically significant according to paired samples T-Test (Table 2). We believe this proves the high degree of familiarization with the SSA system since it indicates users were moving fast to the

possible destination skipping much of the unimportant parts. Users of the SEBS were a bit more careful in reading and spent more time per movement. What is very interesting nevertheless is that users in the SBS condition spent even more time per movement and listened to the text even more carefully. Combined with the previous findings about time spent per page and the fact that they completed successfully their tasks we can understand that they read less and more important parts of the text than the users of the SSA system.

Table 2. Statistically significant differences between SEBS and other systems ($p<0.05$)

Evaluation criteria	*System compared with S.E.B.S*
Task time	SSA
	SBS
Avg. time in visits per task	SBS
Avg. movement time per task	SBS
	SSA
Keystrokes per task	SBS
Avg. Keystrokes per visits and task	SBS

Number of keystrokes per task. The results of the number of keystrokes per page reinforce the previous conclusions. Users with the SBS used less keystrokes to complete their tasks than any other system. However, the fact that users with SEBS used more keystrokes than the ones with the SSA is slightly concerning and indicates that users with SEBS might have used some time from the information seeking tasks to familiarize and explore the web pages visited or might have suffered from disorientation that lead them to listen the same parts of pages many times.

Average number of keystrokes pressed per visit for every task. Findings in this metric follow the results of the previous metric confirming that users with SBS performed better than with any other system. They also reinforce our belief that the familiarization with the SSA system provided enough leverage in comparison with the new and unknown SEBS system.

Text re-entries. Finally the findings in the text re-entries metrics confirmed our doubts that users using SEBS could suffer in many cases from disorientation. In fact we can see that users in the SEBS condition had almost double text re-entries than those with the SSA system. In addition users with SBS had almost half re-entries of users with SSA and only a quarter of those with SBS system.

Discussion of the results. To sum up, the findings from descriptions using SSA reinforced our belief that information given to users through SBS and SEBS about the existence of various elements in web pages can not be perceived and used efficiently in SSA. Moreover the information given to them through SEBS is easily understood and introduces them to web page design which could prove to be helpful in the future.

In overall the SEBS condition, increased the time needed for task completion and the keystrokes pressed. After a second reading of the results though, we can see that this can be explained firstly by the disorientation that seemed to cause. The other

factor is that they were using the spatial information for the first time ever and did not have any knowledge of web page design principles such as the desktop metaphor. Thus, they did not have enough time to familiarize, absorb and learn how to use it efficiently. In fact, many users commented on it, saying that the information gained from SEBS would certainly be more helpful in future visits to web sites. On the other hand, their familiarization with SSA system worked as leverage and reduced the time and key-striking effort when using the SSA. In fact, most of the users have developed strategies such as reading the links of a page or reading a page backwards in order to eliminate the time and key-striking effort needed to reach their destination. The later explains also why they performed better with the SBS system. SBS is much closer to the already used strategies of browsing through links of headings in a web page.

5 Conclusions and Further Research

Summing up, in this paper we presented an application framework based on the semantic web that can provide a basis for improving Web accessibility. Moreover, we presented how the framework utilizes spatial information kept in the annotations through the SEBS feature. To investigate its effects on web browsing and web information seeking we conducted an experiment with 15 blind users that used SeEBrowser under three different conditions (namely SSA, SBS, SEBS).

The analysis of the data showed that the SEBS was not used as expected and did not improve the browsing efficiency. On the other hand, the SBS mechanism improved considerably browsing efficiency. To investigate whether these results can be explained by the fact that users were not familiar with the new feature and spatial information the next experiment will be conducted after a sufficient period of use of SeEBrowser and its features. However, the descriptions of web pages showed that the information provided by the SBS and its variations is quite important. In particular, users tend to miss elements in web pages that depend heavily on their visual representation such as menus and advertisements.

The SEBS feature aimed at reducing the initial list of elements of SBS to a more focused and limited one. Apart from the location of elements there is also another to achieve this. Having as feedback the use of browsing shortcuts in general we can design a learning adaptation mechanism. Up to now the order of the elements in the list is random. However, the feedback from their use can give us information on which shortcuts are preferred more in specific web pages and in general. This information will enable the design of a personalized adaptive browsing shortcuts feature that could reorder and reduce the initial list of annotated elements.

Finally as already pointed out in the framework description there is a vast area of research that could be done on automatic and semi-automatic annotation. Ramakrisnah [8] has already suggested mechanisms to achieve this but they work well only with web pages on a specific domain. The two most important approaches are based on structural analysis of web pages and by applying pattern recognition mechanisms and the other by lexical analysis applying natural language process mechanisms. When combined, these two approaches can give even better results.

Acknowledgments. The research project is funded by the E.U. and the Greek Ministry of Education under the research program "Archimedes" [2].

References

1. Annotea project (2006) http://www.w3.org/2001/Annotea/
2. SeEBrowser Project (Archimedes) (2006) http://seebrowser.it.teithe.gr/
3. Goble, C., Harper, S., Stevens, R.: The travails of visually impaired web travelers. In: Proceedings of the eleventh ACM on Hypertext and hypermedia, pp. 1–10 (2000)
4. Hanson, V., Richards, J.: Achieving a more usable WorldWide Web. Behaviour and Information Technology 24, 231–246 (2005)
5. Huang, A., Sundaresan, N.: A semantic transcoding system to adapt Web services for users with disabilities. In: Proceedings of the fourth international ACM conference on Assistive technologies, pp. 156–163 (2000)
6. Kouroupetroglou, C., Salampasis, M., Manitsaris, A.: A Semantic-web based framework for developing applications to improve accessibility in the WWW. In: W4A Proceedings of the 2006 international cross-disciplinary workshop on Web accessibility (W4A) pp. 98–108 (2006)
7. Marshall, C., Shipman, F.: Which semantic web? In: HYPERTEXT '03 Proceedings of the fourteenth ACM conference on Hypertext and hypermedia, pp. 57–66 (2003)
8. Ramakrishnan, I., Stent, A., Yang, G.: Hearsay: enabling audio browsing on hypertext content. In: Proceedings of the 13th international conference on World Wide Web, pp. 80-89 (2004)
9. Salampasis, M., Kouroupetroglou, C., Manitsaris, A.: Semantically enhanced browsing for blind people in the WWW. HYPERTEXT '05. In: Proceedings of the sixteenth ACM conference on Hypertext and hypermedia, pp. 32–34 (2005)
10. Takagi, H., Asakawa, C., Fukuda, K., Maeda, J.: Site-wide annotation: reconstructing existing pages to be accessible. In: Proceedings of the fifth international ACM conference on Assistive technologies, pp. 81–88 (2002)
11. Tektonidis, D., Bokma, A., Oatley, G., Salampasis, M.: ONAR: An ontologies-based service oriented application integration framework. In: 1st International Conference on Interoperability of Enterprise Software and Applications (2005)
12. Xydas, G., Kouroupetroglou, G.: The DEMOSTHeNES Speech Composer. In: 4th ISCA Tutorial and Workshop on Speech Synthesis (SSW4), pp. 167–172 (2005)
13. Yesilada, Y., Stevens, R., Goble, C.: A Foundation for Tool Based Mobility Support for Visually Impaired Web Users. Paper presented at the Proceedings of the twelfth international conference on World Wide Web (2003)

Chatting with an AAC(Augmentative/Alternative Communication Device)

Scott H. Kuebler and Ronald C. Kuebler

Lingual Perfection, 6601 Eastbrook Road, Columbia, South Carolina, USA Phone: 803.4296730, Fax: 803.7876889,
harrisku@sc.rr.com, lingualperfection@usa.net,
www.lingualperfection.com

Abstract. Communicating with someone using an electronic communication device is different from routine conversation. Despite enormous technological progress, talking with a person with a disability using an electronic communicater is an acquired skill. Very few people in the general population and few professionals know how to chat with someone using an assistive device. Experience a conversation now so you will not have to avoid it when you need to talk to someone who communicates with a device. Let's overcome the fear of talking with someone who uses an AAC device to facilitate communication.

Keywords: Augmentative/Alternative Communication(AAC), Assistive Technology, Down's syndrome, mental retardation, electronic conversation, Picture Communication Symbols(PCS), speech therapy, ChatPC, conversational dynamics, volunteers.

The notion that chatting with an AAC device is different than chatting without one is a challenge to be overcome. Today, you can help by participating in this session to do just that. Someday, people will find it natural to speak with someone who uses assistive technology.

Hello, my name is Ron Kuebler and I am a speech pathologist and operate a private speech therapy practice called Lingual Perfection. Scott and I have a mission to help a person feel comfortable when they talk to someone who uses an assistive technology device. I would like to introduce you to Scott Kuebler, my son, and the person you will be talking with today. Scott is using a ChatPC by Dave Hershberger and Saltillo Corporation and it is, basically, a Personal Desk Assistant(PDA) with an external amplifier. It uses software incorporating Picture Communication Symbols(PCS) by Mayer-Johnson and digital images and text with synthetic voicing. Let's hear from Scott.

Hello, my name is Scott Kuebler(pictured)… Scott uses the ChatPC to report on his personal details. He will also talk about his employment and travel experiences in promoting speech with assistive technology. Scott is a 26 year old man with Down's syndrome and uses the ChatPC to clarify what he says. His use of Augmentative/Alternative Communication(AAC) devices goes back to 1996 when was 16 years old and he used the Pegasus Lite by Words+. He has an expressive/receptive language disorder secondary to Down's syndrome and has received speech therapy since he

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 368–370, 2007.
© Springer-Verlag Berlin Heidelberg 2007

was a young boy. Before using electronic devices, he used manual communication boards with marginal success.

Scott Kuebler lives at home with his parents, Ron and Margot Kuebler and works at the BiLo grocery store for a few hours each week as a bagger. It is within one mile(1.6 kilometers) and he walks back and forth to work. Scott has been a member of the Boy Scouts of America for many years as a boy and as an adult leader and is an Eagle Scout, the highest rank in scouting. Scott also earned the Gold Congressional Award in 2000 and promotes this program around the world as the International Award and by other names. Scott has two brothers, Gregory(24) and Matthew(22) Kuebler, who live in other parts of the United States of America.

How may people have engaged in conversation with someone using an assistive technology device is unknown to the authors. However, in the experience of presenting to individuals involved in the field of AAC, the numbers are surprisingly small. When the general public is requested to indicate experience, the results observed by the authors is dismally small.

Conversational dynamics are considerably different when one of the communication partners uses an AAC device. The major issue is the fact that most people do not know what the communication device does or how it is accomplished. Being empathic certainly helps as most AAC users or their caregivers are able to provide guidance as to how a conversation should go.

A general conversation may have different dynamics depending on the cultural and linguistic environment. This presentation introduces a Standard English context, although the authors are interested in hearing about variations presented by audience members. Certainly the comparison of general conversation with and without an AAC device is affected by the cultural and linguistic parameters.

Dialectic variations certainly cause confusion in a conversation when listeners are unfamiliar with the words or phrases. For example, an American from the South of the United States immediately knows what you mean when you say, “ya’ll”. It means that a person is addressing everyone within listening distance and could be one or more people. Rate of presentation of a thought is affected when using an AAC device because the process time is extended. Rate of reception may be extended as well for the transmitter as well as the receiver. Intelligibility is affected by the use of picture and digital images to transmit a thought. Person listening and seeing something transmitted by an AAC user may ask for increased repetitions to clarify the item. An AAC user may have increased repetition requests as well, especially if the non-AAC user is not empathic to the user’s method of communication. When we speak, we can control our loudness spontaneously, but an AAC user may need to turn dials to adjust volume for the situation and noise level. Some individuals may have difficulty interpreting symbols used by the AAC user. The time to formulate a thought when using an AAC device is extended and dependent on the competence of the individual to access information and to transmit it. What other factors can anyone think of?

A structured conversation will be developed during the presentation to give audience members an increased understanding of what they might want to say and ask in a volunteer conversation. This also highlights the methods that can be used to negotiate the actual conversation. Facial expression can tell either communication partner how comfortable each is with a conversation. Body language also indicates a measure of comfort during an augmented conversation. Sometimes people will change the tone

and other features of their voice when speaking to someone using an AAC device. Beginning and ending a conversation may consist of similar words when comparing an augmented and non-augmented conversation. One extra step is working through the device to send and receive the message.

Volunteers will be asked to engage Scott in a conversation and the audience will be given opportunity after the conversation to make comments and suggestions. Subsequent conversationalists may benefit from their predecessors as to the dynamics of the subsequent conversations.

The importance of talking about the conversations has immediate and long range effects. Audience members will be challenged to take their new or improved skills back to the community in which they live and engage people who use AAC. The other expectation would be to share the experience with other individuals who would like to learn how to communicate with someone who uses assistive technology.

Advances in the area of AAC technology will continue in miniaturization of devices, scope of control by one device, communication prediction or completion, and speed tasks are carried out. Unfortunately, the rate of assimilation of advanced techniques by users will probably lag behind due to the lack of coordinated teaching of AAC use to users and their mentors. A long range plan is needed to counter this ongoing trend.

References

1. Kuebler, S.H., Kuebler, R.: Promoting an Idea with AAC. In: Proceedings of ISAAC 2004, Natal, Brazil (2004)
2. Kuebler, R., Kuebler, S.H: Dealing with Grief with AAC. In: ATIA Conference Proceedings, Orlando, Florida, USA (2005)
3. Kuebler, R., Kuebler, S.: Augmented Conversations. In: ABDA/ABMPP Proceedings, Chicago. Illinois, USA (2005)
4. Kuebler, S.H., Kuebler, R.: AAC Role Play – Restaurant Order. In: Proceedings of ISAAC 2006, Dusseldorf, Germany (2006)
5. Kuebler, R., Kuebler, S.: AAC in Action – At Work. In: ATIA 2007, Conference Proceedings, Orlando, Florida, USA (2007)

A Flexible On-Screen Keyboard: Dynamically Adapting for Individuals' Needs

Yun-Lung Lin[1], Ming-Chung Chen[2], Ya-Ping Wu[3], Yao-Ming Yeh[1], and Hwa-Pey Wang[3]

[1] Department of Information and Computer Education, National Taiwan Normal University
162 Ho-Ping E. Rd., 1 Sec., Taipei, Taiwan
{harrison,ymyeh}@ice.ntnu.edu.tw
[2] Department of Special Education, National Chiayi University
300 University Road, Chiayi, Taiwan
mtchen@mail.ncyu.edu.tw
[3]Department of Special Education, National Taiwan Normal University
162 Ho-Ping E. Rd., 1 Sec., Taipei, Taiwan
{893090036,e14008}@.ntnu.edu.tw

Abstract. This study was to design an alternative on-screen keyboard and evaluate the efficacy of this innovative layout design for people with severe physical disability. The matrix keyboard layouts was designed based on human computer interaction. A repeated experiment was performed to compare the speed and accuracy of text entry with point-and-click input method between the matrix on-screen keyboard and the Windows XP QWERTY virtual keyboard. Data analysis revealed that the matrix on-screen keyboard provided better performance for the participant. The result also indicates that layout adaptation assessment is a valid tool to confirm proper layout size for users and that alphabetic order is better than QWERTY order for a novice user to learn a new on-screen keyboard. A usability study was undertaken to evaluation the performance of the double click instead of the point-and-click plus Shift key input method. The possible causes of the results and suggestions for further studies are discussed.

Keywords: accessibility, on-screen keyboard, scanning selection.

1 Introduction

Computers allow us to communicate, work, learn, and perform many leisure activities; they also hold great potential for influencing the lifestyles of people with disabilities. Using a standard keyboard might seem quite normal to most people, but using it can be a strenuous task for people with motor disabilities. There are many alternative on-screen keyboards available to people who have difficulty using the standard keyboard. The purpose of this study was to design an alternative on-screen keyboard to improve the entry efficiency of people with motor disabilities. The efficiency of on-screen keyboards depends on the sustainability of the control interface, including proper selection methods and keyboard layouts [1]. Designing for

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 371–379, 2007.
© Springer-Verlag Berlin Heidelberg 2007

diversity was a factor we considered carefully in this study [2]. The keyboard interface design for people with motor disabilities requires the consideration of many factors reflecting the diversity of disability types and disability degree. We have implemented a one handed numeric-based input method for people with mild motor disabilities and automatic scanning for people with severe motor disabilities. To increase the performance of text entry, we also implemented a new approach to the point-and-click input method.

2 On-Screen Keyboards

The keyboard is the most common method of controlling the computer. The drawbacks of the QWERTY keyboard are in its two-handed design and its requirement of the full use of all ten fingers. Thus, people with physical disabilities have difficulties in accessing a standard keyboard. They can interaction with computers by using on-screen keyboards and assisted peripherals. The purpose of assessing physical motor control is to find the most effective input device and selection technique. Lee and Thomas describe the various components of assistive technology access including input device, selection set, selection technique, and application information [3]. Text entry efficiency is dependent on an alternative interface and on the quality of interaction between the user and the access method.

2.1 Selection Technique

Most of the on-screen keyboard support point-and-click, dwell, and scanning input methods. Point-and-click access is suitable for people with limited movement range who cannot reach across the standard keyboard, but can operate a pointing device such as a trackball or touch pad. If the user is unable to press a button or switch, but is able to keep the pointing device steady for a short time, dwell selection is an alternative solution. For users with severe physical limitations who can physically control only one or a small number of muscle movements, scanning is the only viable control option.

Switch-based scanning is typically used by persons with severe physical impairments. These impairments may be due to disorders such as cerebral palsy, acute muscle weakness at advanced stages of muscular dystrophy, or paralysis due to high spinal cord injury. Scanning involves the successive presentation of items in a display from which the user selects a desired item by activating a switch when the item is highlighted. Three types of scanning with a single switch are distinguished: automatic, step, and inverse. Single-switch automatic row-column scanning is the most common scanning approach.

2.2 Selection Technique

Keyboard layout is an important variable in affecting typing speed, typing accuracy, and ease of learning. Typically, on-screen keyboards display the QWERTY keyboard; however, the QWERTY order layout can be confusing for novice users. In addition to QWERTY, layout character sets can be arranged according to alphabetic order or frequency of use. The frequency of use keyboard groups commonly typed characters

to reduce physical effort and help increase typing speed. Although the frequency layout is theoretically physically fastest, input on the keyboard layout with alphabetical ordering will initially be easier for novice users.

2.3 Features to Improve Performance

Several commercial on-screen keyboards, such as WiVik, Grids, WinScan and Click-N-Type, provide more sophisticated functions to allow the improved typing performance of people with disabilities. These programs support computer access features like macros, menus, and dialog boxes; some on-screen keyboards support acceleration features. A common approach to enhancing text input rates is to use word prediction. A number of word prediction systems typically offer users a list of suggested completions of the current word. Users can choose a word from the list using a single command. Other acceleration features are abbreviation expansion and smart punctuation. In addition to controlling text entry, many on-screen keyboards have other features to improve ease of use, such as automatic start up, saving and loading user settings, control over applications, and sound actions[4] [5].

3 Design and Method

3.1 Chorded Input Method

We implemented a chorded input method that allows the user to generate a character by pressing two numeric keys. This chorded input method can be operated with a signal hand. As shown in Figure 1, the first key press selects a sub-matrix, and the second key press selects a letter or symbol in that sub-matrix. For example, when using the numeric keypad, the user presses '4' and then '7' to output the corresponding letter 'a' to the application program. Pressing '5'and '5' will generate the space character. The user can use '0' as a cancel button when a typing error occurs. Based on Fitts' law of movement efficiency, punctuation and commands were placed according to their frequency of use to reduce the fingers' travelling distance.

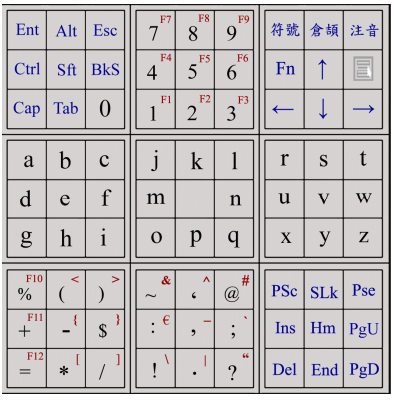


Fig. 1. Keyboard layout design

Fig. 2. Pop-up menu

The goal of universal design is to make objects accessible to everyone. Based on the aim of designing for a diversity of users, our on-screen keyboard supports a wide range of access methods and adjustments. We also added a high contrast layout for visual impaired users. The on-screen keyboard is displayed within a window that can be moved, resized and supports audio feedback to the user. As show in Figure 2, it also includes a pop-up menu to allow for reconfiguration by a right click of the on-screen keyboard.

3.2 Scanning Input Method

Group-row-column scanning is typically more efficient than row-column scanning. The Microsoft Windows XP operating system and some commercial on-screen keyboard applications support group-row-column scanning with the QWERTY layout. We developed a method of group-row-column scanning based on our matrix layout. The major matrix is divided into 9 sub-matrices. During initial input, the user is expected to select the sub-matrix where the target symbol belongs. When the highlight arrives at the desired group, the user presses a switch. The highlight then advance scans each row in the group until the user presses the switch again. Finally, the highlight scans across each item in the selected row until the user presses the switch to select the character. Figure 3 shows 9 groups of the matrix on-screen keyboard and the sequence of group scanning is 4, 5, 6, 7, 8, 9, 1, 2, 3.

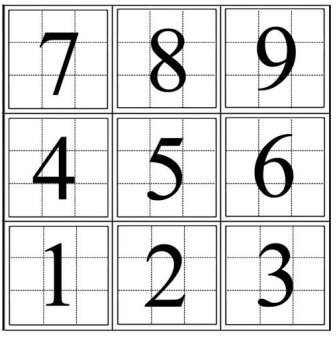


Fig. 3. 9 groups of scanning

The order of the symbols in the scanning keyboard is crucial for the keyboard's performance, since it constrains the access time of each character. We took into consideration letter frequencies in designing our keyboard layout. Figure 4 illustrates the number of scan steps needed to select each item in our matrix. An important principle guiding our design was the notion that different letters occur at different rates in a body of text. Previous research has shown that the letter e occurs about 13% of the time, while the letter z appears at a rate of about 0.2% in a body of English text**[6]**. The letters in the scanning layout are arranged according to frequency of use, which is shown in Figure 5. We also extended the keyboard to special characters such as punctuation, function keys and numeric keys. Since the space character is

6	7	8	7	8	9	8	9	10
7	8	9	8	9	10	9	10	11
8	9	10	9	10	11	10	11	12
3	4	5	4	5	6	5	6	7
4	5	6	5	6	7	6	7	8
5	6	7	6	7	8	7	8	9
9	10	11	10	11	12	11	12	13
10	11	12	11	12	13	12	13	14
11	12	13	12	13	14	13	14	15

Fig. 4. Switch counts in Matrix layout

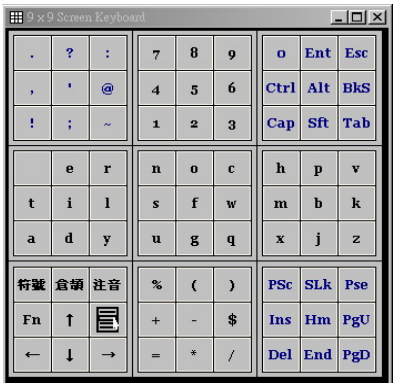


Fig. 5. Scanning layout design

prominent in text entry, it is located on the highest priority position. It was expected that the rearrangement of the keyboard layout according to frequency of use would improve the typing rate by 30-50 % [7].

3.3 Point-and-Click Input Method

Visually based point-and-click interactions are very common for on-screen keyboards. Most on-screen keyboards require use of the Shift key to type the upper case and some punctuation. To minimize the distance of movement, we developed an innovative text entry technique for people with motor disabilities. We designed our method to use the double click instead of the shift key. In our system the user can employ a single-click to output a lowercase letter or number, and a double click to output an uppercase letter, function key, or punctuation.

4 Study Methodology

4.1 Participant

To evaluate the usability of the point-and-click input method, text entry tasks were performed to measure the typing performance of a participant. The participant has a severe physical impairment (Osteogenesis imperfecta). Although she had used computers in the past for three years but she was not familiar with the QWERTY keyboard layout; she had used the point-and-click mode with her left hand. Her performance was poor when using the virtual keyboard embedded in the Windows XP operating system.

4.2 Materials and Apparatus

The evaluation tools in this study included two computer software programs that were designed for this experiment, "Layout Adaptation Assessment" and "Entry Performance Evaluation".

4.2.1 Layout Adaptation Assessment

Layout size is a very important element when considering movement efficiency. A full selection set with large items is obviously easier to access with the point-and-click input method. However, such size requires a considerable amount of space on the screen and increases the travel distance of the pointer. A smaller layout will reduce movement time requirements, but increase the typing error rate. Thus, travel distance and accuracy are the key considerations of when designing for text entry. Our on-screen keyboard supports the re-sizing of the layout and automatically adjusts the size of the individual items correspondingly. Proper layout size is selected for the user by the use of an evaluation tool; our program has seven preset layout sizes. Adjusting the on-screen keyboard to an appropriate layout size and using a double click instead of the shift key should give significant performance improvements for text entry.

The "Layout Adaptation Assessment" program was used to determine the appropriate keyboard layout for the participant. A picture of the program is shown in Figure 6. We applied a binary search tree algorithm to select the appropriate layout for the participant, which is shown in Figure 7. The user selected the proper layout from seven different layout sizes. LAA based on a binary search tree of size 7 and depth 3. A set of three typing tasks was required to be completed for each layout. To counterbalance the effect of learning, the sequences of three tasks were assigned randomly in the assessment procedures.

Task 1: x, (, 1, b, (, x, 1, (, b, 1, x, b, x
Task 2: n, o, k, l, o, n, k, o, l, k, n, l, n
Task 3: g, k, 5, 9, k, g, 5, k, 9, 5, g, 9, g

During the initial assessment of each user the default layout size is type 4 (7.5cm x 7.5cm). A smaller layout is available when the initial assessment result reaches the criterion of 90% accuracy. A larger layout will be allocated when the user cannot reach this criterion. The same procedures will be repeated until the proper layout size for the user is found.

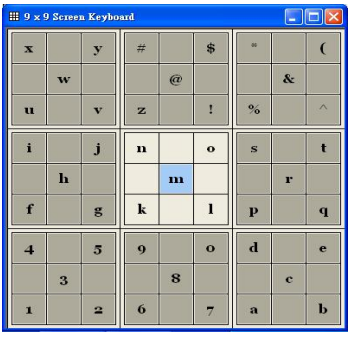


Fig. 6. Layout of adaptation assessment

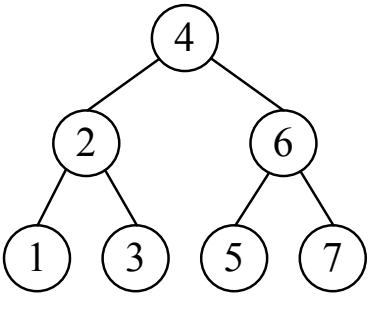


Fig. 7. Layout select algorithm

4.2.2 Entry Performance Evaluation

During our testing, an alphabetical character or punctuation was generated randomly and shown on the upper square of the screen. Some characters required the use of the Shift key in conjunction with other keys. The participant was asked to type the indicated character on our on-screen keyboard and the Windows XP virtual keyboard. A picture of the test application is shown in Figure 8.

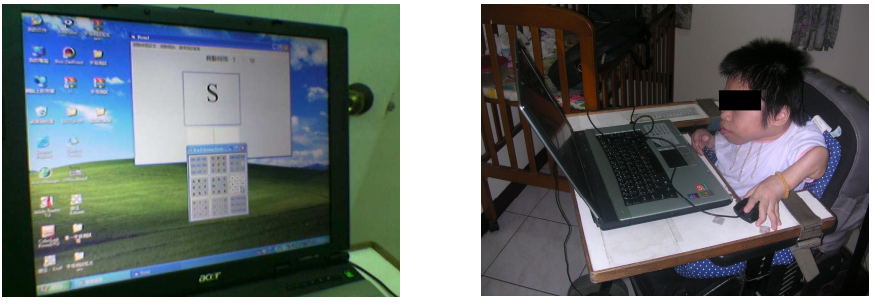


Fig. 8. Typing performance test of Point-and-click Input Method

4.3 Procedure

A repeated experiment was performed to compare the typing speed and accuracy between the two different keyboard layouts. The experiment was conducted with a 15-inch laptop computer in the participant's home. The experiment started with the researcher explaining the task and using the "Layout Adaptation Assessment" program to determine the appropriate keyboard layout for the subject. The result showed that type 2 (6.5cm x 6.5cm) is the proper layout for the participant. After establishing the proper size, the participant took 20 minutes to practice the matrix on-screen keyboard and Windows on-screen QWERTY keyboard.

A picture of the evaluation situation is shown in Figure. 8. A character was displayed on screen and the participant was asked to type the characters as accurately as possible, without performing any error correction. Keyboard trials were conducted twice a day over a period of five days.

Each session was administered in the same way by the researcher. This involved adjusting the table height and positioning the monitor and mouse appropriately. Each session had a dissimilar sequence of two different keyboard layouts to counterbalance the effect of learning. This process was repeated until the subject finished all the test session of the two different keyboard layouts.

5 Result and Discussion

Text entry speed and accuracy are often used as performance measures for evaluating input systems. As shown in Figure 9, we obtained fifteen measurements of text entry speed. Using the matrix on-screen keyboard, the participant could type only 28 correct characters in the first session, but was eventually up to 103 characters in the

last session; whereas the results of the QWERTY layout of the Windows XP on-screen keyboard started at 24 characters and ended at 94 characters. The result indicates that the participant performed better while using the matrix on-screen keyboard than the using Windows XP on-screen keyboard.

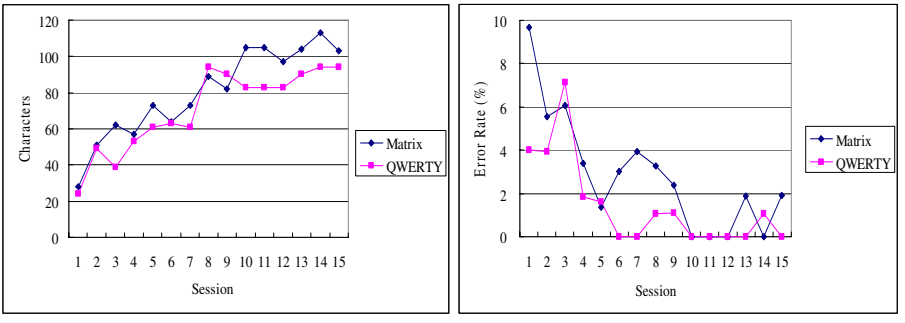


Fig. 9. Entry speed and error rates by variant and session

The participant occasionally made errors, there was no significant difference between the error rates on each keyboard layouts. We also wanted to know how the participant felt about the new approach. An interview was conducted with the participant after the final session. The participant indicated that "Layout Adaptation Assessment" is a valid tool to confirm the proper layout size for users. She also stated that the alphabetic order is better than QWERTY order for a novice user.

6 Conclusions

The need for participation in an information society has let to several research efforts for designing accessibility solutions for people with disabilities. People with motor disabilities do not have the ability to efficiently use standard input devices. This paper has introduced a new on-screen keyboard for use in combination with selected layout sizes. The experiments demonstrate the effects of our on-screen keyboard in comparison to the Windows XP virtual keyboard. The result reveals that the matrix on-screen keyboard worked more effectively than the QWERTY keyboard in the usability evaluation. However, during the evaluation, each input method session only lasted 5 minutes. User fatigue may be a potentially problematic variable that must be evaluated in future studies.

A usability study was undertaken to evaluate the performance of the double click in comparison to the point-and-click plus Shift key input method. Future research that validates the effectiveness of other access methods for people with motor disabilities is needed.

Speed of entry is not the only criterion for an efficient input system. We should take into account other issues such as ease of learning, low error rates, and simple error correction. The time spent on correcting typing errors has a major impact on text

entry efficiency. A potential limitation of this study is its reliance on only one participant, which may limit the research results to the people with motor disabilities.

To confirm with the concept of dynamic diversity, on-screen keyboards need to be suited to the changing requirements of the user. Successful computer access often requires a combined approach, an on-screen keyboard will rarely provide a complete solution. If alternative computer input devices are necessary, it is likely that the hardware settings will need customizing. People with motor disabilities may need to adjust the configuration of their input system, but often find it is an obscure and difficult process. By means of a number of the evaluation sessions, the system ideally could, according to the user's preferences and abilities, automatically be self-customizing.

The success of input systems depends on numerous physical and psychological factors. As for future work, we will advance our system by making it compatible with output adaptive tools, to provide better solutions for people with motor disabilities.

Acknowledgements. The authors thank the participant involved this study. The first author expressed his gratitude to his mentor, Dr. Tein-Yu Li, who kept inspiring him to finish this study. We also would like to thank Miss Ching-Hui Wang for her effort in coding this system.

References

1. Lehikoinen, J., Salminen, I.: An empirical and theoretical evaluation of BinScroll: a rapid selection technique for alphanumeric lists. Personal and Ubiquitous Computing, pp. 141–150 (2002)
2. Lopes, J.B.: Designing user interfaces for severely handicapped persons. In: Workshop on Universal Accessibility of Ubiquitous Computing, pp. 100–106 (2001)
3. Lee, k. S., Tomas, D. J.: Control of computer-based technology for people with physical disabilities: An Assessment Manual. University of Toronto Press (1990)
4. Cook, A.M., Hussey, S.M. (eds.): Assistive technology: Principles and practice (2nd edn.) Baltimore: Mosby (1990)
5. Istance, H.O.I., Spinner, C., Howarth, P.A.: Providing Motor-Impaired users with Access to Standard Graphical User Interfaces (GUI) Software via Eye-Based Interaction. In: Proceedings of the1st European Conference on Disability, Virtual Reality and Associated Technologies (ECDVRAT'96), Maidenhead, UK, pp. 109–116 (1996)
6. Nesbat, S.B.: A system for fast, full-text entry for small electronic devices. In: Proceedings of the 5th international conference on Multimodal interfaces, Canada, pp. 4–11 (2003)
7. Steriadis, C.E., Constantinou, P.: Designing Human-Computer Interfaces for Quadriplegic People. ACM Transactions on Computer-Human Interaction 10, 87–118 (2003)

Evaluation of RSVP and Display Types on Decoding Performance of Information Extraction Tasks

Ya-Li Lin and Chun-Min Ho

Department of Statistics, Tunghai University, 40704Taichung, Taiwan
{Ya-Li Lin,Chun-Min Ho,yllin}@thu.edu.tw

Abstract. The decoding performance of information extraction is important for data analysis and decision making. The objective of this study is to evaluate the effects of display types (TYPE) shown in web pages using different techniques of rapid serial visual presentation (RSVP) on decoding performance for different information extraction tasks (TASK). Twenty-four subjects (12 males, 12 females) participated in the information extraction tasks. Two factorial designs with blocking experiment were used to collect the response time data. In Experiment 1 (TASK includes finding exact value and identifying the trend), the results showed that both the 2-way interactions of RSVP and TASK as well as TYPE and TASK would significantly affect mean response time. In Experiment 2 (TASK includes finding maximum and comparing two values), the results showed that the 3-way interaction of RSVP, TYPE, and TASK would significantly affect mean response time.

Keywords: Rapid Serial Visual Presentation, Dynamic Visual Interfaces, Graphical Perception, Quantitative Information Extraction, Decoding Performance.

1 Introduction

The decoding performance of quantitative information extraction is important for data analysis and decision making. Statistical graphs and alphanumerical displays of quantitative data are major components in software applications, e.g., statistical packages, spreadsheets, and decision support systems [3]. Meyer proposed a minimum assumption Visual Search Model (VSM) to predict the performance of various tasks using tables and graphs with different levels of complexity of data. The VSM predicted the results for tables well [2]. For graphs the model was adequate for tasks that involved single data points, such as reading values or comparing pairs of values. The performance of tasks that required the analysis of data configurations, such as reading a trend, could not be predicted with the VSM.

Compared with the displays of static graphics, static tables, and static texts, rapid serial visual presentation (RSVP) was used as the electronic equivalent of riffling a book [1]. Moreover, applications of RSVP might generally provide useful tools for the presentation of large numbers of loosely structured alternatives that could be adequately represented by image preview [4], [5].

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 380–388, 2007.
© Springer-Verlag Berlin Heidelberg 2007

We concentrated on the impact of design decisions underlying a number of existing modes of RSVP. The images of conventional mode and keyhole RSVP mode shown in the fixation area are used to compare with carousel, collage, and shelf RSVP modes. Four display types (abbreviated as TYPE) including bar charts, line graphs, tables, and texts are used to present information extraction tasks. The objective of this study is to evaluate the effects of display types (TYPE) shown in the web pages using different techniques of RSVP on decoding performance for different information extraction tasks (TASK). In Experiment 1, TASK includes finding exact value and identifying the trend and in Experiment 2, TASK includes finding maximum and comparing two values. The researchers' hypotheses will be discussed in both Experiment 1 and Experiment 2: Do the effects of RSVP modes, display types, and TASKS interact to influence the visual decoding performance? Two factorial designs with blocking experiment including the combinations of four image display types, five RSVP modes, and two tasks will be used to collect the response time data.

2 Methods

2.1 Participants

Twenty-four subjects (12 males, 12 females) participated in the information extraction tasks with mean age of 22.6 years old and standard deviation of 2.2 years old. In Experiment 1, twelve subjects (6 males, 6 females) were assigned to two kinds of information extraction tasks: reading exact values (Task 1) and identifying trends (Task 2). In Experiment 2, another twelve subjects (6 males, 6 females) were assigned to two kinds of information extraction tasks: finding maximum values (Task 3) and comparing values (Task 4). All subjects have to fulfill a course requirement in an introductory Statistics course. Each participant was paid after experiment.

2.2 Apparatus and Stimuli

The experiment used a Pentium IV desktop computer (256MB RAM) with a 17-inches CRT screen (1280×1024 pixels) and a web camera (Logitech's QuickCam Pro5000) with a built-in microphone. Microsoft Internet Explore 6.0, Macromedia Dreamweaver 4.0 (2000) and Ulead Graphic Interchange Format (GIF) Animator 4.0 would be used to design the webpage. In addition, CyberLink StreamAuthor 3.0 was used to record the process of operation displays. A digital video camera recorder (SONY DCR-PC330) was used to record the overall process of experiments.

The displays presented data as four types of texts, tables, line graphs, and bar graphs (TYPE factor) using five RSVP modes: keyhole, carousel, collage, shelf, and conventional mode (RSVP factor) with user self-control presentation rates. The treatment combinations of RSVP modes and TYPE displays were illustrated in Figures 1 to 4. The contents of questions bank came from the public information announced by Taiwan Stocks Exchange Corporation for the top 10 market capitalization companies. The questions in Experiment 1 are illustrated in Figure 5 which included reading the exact value of a data point (e.g., "How much the closing price was for Foxconn at September 29, 2005?") and identifying a trend in a series (e.g., "Was the general trend of the closing price increasing or decreasing for Formosa

Plastic Corporation from November 30 to December 8, 2005?"). The questions in Experiment 2 are illustrated in Figure 6 which included finding maximum values (e.g., "Choose the maximum closing price from November 23rd to December 2nd, 2005 for Nan Ya Plastics Corporation.") and comparing values (e.g., "How many days the closing price was 23 NTD from June 21 to June 30, 2005 for UMC?").

2.3 Design

The dependent variable was response time in Experiment 1 and Experiment 2. Response times were analyzed via two three-way factorial designs with blocking experiment with display types (texts, tables, line graphs, and bar graphs), RSVP modes (keyhole, carousel, collage, shelf, and convention mode), and TASK (Tasks 1 and 2 in Experiment 1, and Tasks 3 and 4 in Experiment 2). All independent variables were within-subjects variables and the various main effects and interaction effects will be discussed. Each participant responded to 40 trials and the order of trials within a block was individually randomized for each participant.

2.4 Procedure

Participants attended the experiment individually. They sat in front of the CRT at a viewing distance of approximately 50cm facing a microphone that was connected to a web camera with Stream Author 3.0 being activated. The height of the center of the CRT and the microphone were individually adjusted to the height of the participant. First, instructions were presented on the computer screen that explained the experimental tasks and the displays used. This presentation was self-paced. Both in Experiments 1 and 2, twelve participants are shown forty treatment combinations of five RSVP modes, four display types, and two kinds of tasks (480 observations in each experiment).

Fig. 1. RSVP modes bar charts from left to right: convention, keyhole, carousel, collage and shelf

Fig. 2. RSVP modes using text displays from left to right: convention, keyhole, carousel, collage and shelf

Fig. 3. RSVP modes using table displays from left to right: convention, keyhole, carousel, collage and shelf

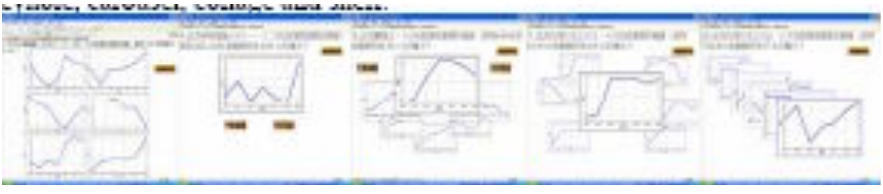


Fig. 4. RSVP modes using line graphs from left to right: convention, keyhole, carousel, collage and shelf

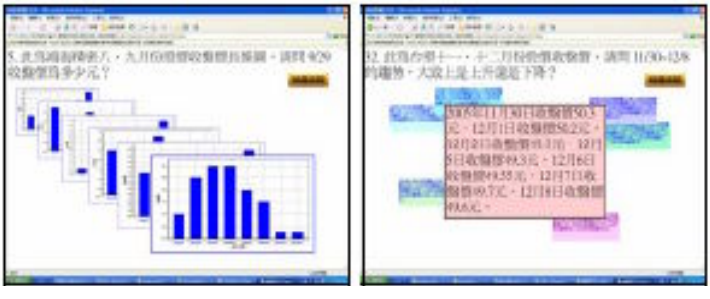


Fig. 5. Illustration of the question for reading an exact value (Task 1) in terms of bar chart and shelf mode (left) and identifying trends (Task 2) in terms of text and collage mode (right)

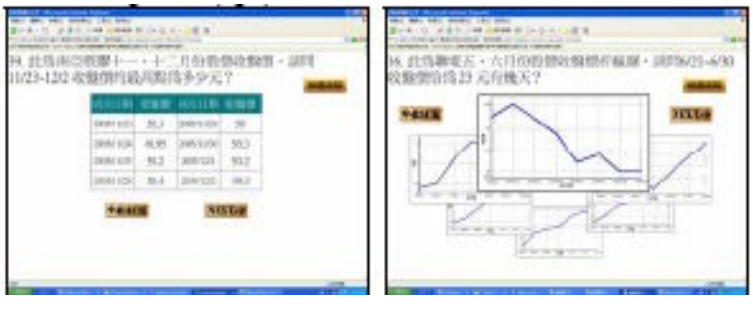


Fig. 6. Illustration of the question for finding maximum values (Task 3) in terms of table and keyhole mode (left) and comparing values (Task 4) in terms of line chart and carousel mode (right).

3 Results and Discussion

In terms of aids of digital video and webcam, the process of executing the experiment and the voices of answering the questions for each participant will be simultaneously recorded with the aids of Stream Author 3.0. Based on the records, the response time will be computed after experiments. The descriptive statistics for the response time in Experiment 1 and Experiment 2 are summarized as Table 1. Each combination has 12 observations. Data in Table 1 indicated that mean response time would be different as the information extraction was visualized on the different display types using different RSVP techniques for different tasks.

In Experiment 1, a three-way ANOVA with blocks model was considered. A plot of residuals against the fitted values showed that error variances are not constant. The Box-Cox procedure suggests a transformation of natural logarithm scale is appropriate for correcting the unequal error variances [6]. The results of ANOVA with log response time were shown in Table 2. The RSVP×TASK interaction ($F(4,429)=2.49$, $p<0.05$) and TYPE×TASK interaction ($F(3,429)=28.48$, $p<0.001$) significantly affected mean log response time. However, there was no enough evidence to provide the statistical significance of RSVP×TYPE×TASK interaction ($F(12,429)=1.18$, $p>0.05$).

For the post hoc pair comparisons of RSVP×TASK interaction, the results of Duncan multiple range tests showed that the better designs of RSVP×TASK are convention/exact value, keyhole/exact value, shelf/exact value, and keyhole/trend combinations and the worse design is collage/trend combination which are shown in left-hand side of Figure 7. For the reading of exact values (Task 1), response times were faster for convention mode than for the other RSVP mode (left-hand side in Figure 8(a)). For the identifying of trends (Task 2), response times were faster for keyhole mode than for the other RSVP mode (right-hand side in Figure 8(a)).

For the post hoc pair comparisons of TYPE×TASK interaction, the results of Duncan multiple range tests showed that the better designs of TYPE×TASK are table/exact value and text/exact value combinations and the worse designs are table/trend, line/table, bar/exact value, and line/exact value combinations which are shown in right-hand side of Figure 8. For the reading of exact values (Task 1), response times were faster for tables and texts than for the other display types (left-hand side in Figure 8(b)). Response times for tables were significantly faster than for bar graphs, which are consistent with the findings of Meyer, et al. [3]. However, for the identifying of trends (Task 2), response times were faster for bars and texts than for other display types (right-hand side in Figure 8(b)). Response times for bars and texts were significantly faster than for line graphs and tables, which reveal different findings with those of Meyer, et al. [3] and Meyer [2].

In Experiment 2, a three-way ANOVA with blocks model was considered again. A plot of residuals against the fitted values also showed that error variances are not constant. The Box-Cox procedure suggests a same transformation of natural logarithm scale is appropriate for correcting the unequal error variances [6]. The results of ANOVA with log response time were shown in Table 2. The RSVP×TYPE×TASK

Table 1. Summary Statistics of response time for four tasks in Experiment 1 and Experiment 2 (unit: second)

TASK		Exact value	Trend	Maximum	Comparison
RSVP	TYPE	Mean (std)	Mean (std)	Mean (std)	Mean (std)
convention	bar	14.48(4.01)	15.16(6.33)	15.96(6.45)	17.44(7.27)
	line	17.11(5.15)	19.02(10.21)	19.74(9.25)	16.56(7.02)
	table	8.31(3.98)	21.47(11.37)	16.18(8.19)	25.71(9.36)
	text	10.43(3.43)	13.74(8.85)	13.54(5.68)	15.72(6.65)
keyhole	bar	16.08(8.25)	9.73(4.09)	16.57(4.42)	15.16(6.98)
	line	18.56(11.1)	20.62(10.92)	15.87(6.49)	15.50(4.90)
	table	12.15(5.92)	18.98(7.80)	14.78(4.55)	17.53(6.46)
	text	10.26(3.03)	11.53(4.15)	14.30(4.46)	16.29(6.59)
carousel	bar	22.29(11.5)	10.65(2.41)	21.62(8.90)	17.83(6.30)
	line	17.51(6.68)	19.92(6.33)	15.86(4.35)	16.75(4.79)
	table	16.29(7.92)	20.87(7.42)	22.24(9.36)	15.58(4.59)
	text	13.51(5.48)	16.31(10.48)	21.89(10.43)	16.30(7.00)
Collage	bar	23.20(10.23)	17.35(6.05)	24.19(11.82)	17.89(6.59)
	line	19.44(5.77)	19.27(13.63)	24.06(12.29)	20.56(6.46)
	table	12.30(5.34)	22.48(10.12)	14.60(4.70)	22.18(6.77)
	text	14.02(4.44)	14.24(5.74)	16.78(4.77)	13.86(5.31)
shelf	bar	19.87(10.23)	15.18(7.35)	17.35(8.89)	15.43(7.19)
	line	19.53(7.88)	18.28(8.90)	16.85(6.42)	19.68(10.60)
	table	10.02(4.93)	18.83(7.16)	14.55(6.05)	19.43(7.98)
	text	10.93(3.38)	16.27(8.65)	16.04(3.08)	12.41(4.04)

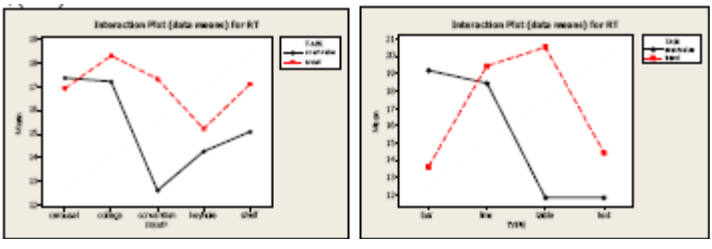


Fig. 7. Interaction plots of mean response time for the five RSVP (l*eft*) and the four display types (*right*) of two tasks in Experiment 1 (unit: second)

interaction ($F(12,429)=2.01$, $p<0.05$), RSVP×TASK interaction ($F(4,429)=2.71$, $p<0.05$) and TYPE×TASK interaction ($F(3,429)=5.68$, $p<0.001$) significantly affected mean log response time.

For the post hoc pair comparisons of RSVP×TYPE×TASK interaction, the results of Duncan multiple range tests showed that, for the finding of maximum values

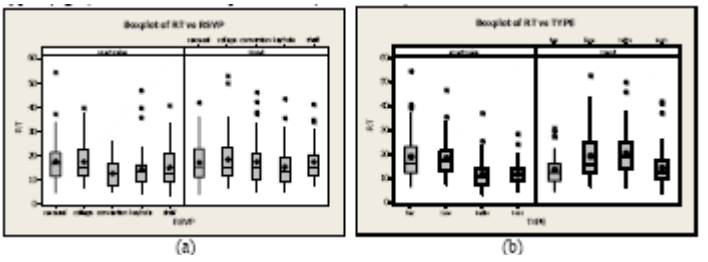


Fig. 8. \ Response time for (a) the five RSVP (*left)* and (b) the four display types (*right)* of two tasks in Experiment 1 (unit: second) where " ⊕ " denotes the group mean

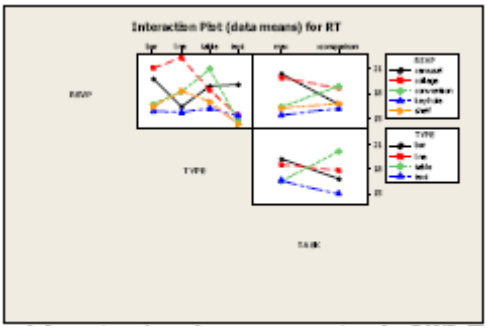


Fig. 9. Interaction plots of mean response time for RSVP×TYPE×TASK in Experiment 2 (unit: second)

Table 2. F-values of three-way analysis of variance with blocks after log transformation

Source of variation	DF	Sum Squares	Mean Square	*F*-value in Experiment 1	Sum Squares	Mean Square	*F*-value in Experiment 2
Block	11	18.22	1.66	**10.94*** **	16.92	1.54	**12.60*** **
RSVP	4	3.41	0.85	**5.64*** **	2.37	0.59	**4.86*** **
TYPE	3	8.18	2.73	**18.02*** **	1.57	0.52	**4.29** **
RSVP* TYPE	12	2.80	0.23	1.54	2.16	0.18	1.48
TASK	1	1.26	1.26	**8.32** **	0.01	0.01	0.09
RSVP* TASK	4	1.51	0.38	**2.49***	1.32	0.33	**2.71***
TYPE* TASK	3	12.93	4.31	**28.48*** **	2.08	0.69	**5.68*** **
RSVP TYPE* TASK	12	2.15	0.18	1.18	2.94	0.25	**2.01***
Error	429	64.95	0.15		52.38	0.12	
Corrected Total	479	115.42			81.77		

(Task 3), the better combinations of RSVP and TYPE are convention/text, shelf/table, keyhole/text, and collage/table, and the worse combinations are collage/bar, collage/line, carousel/table, carousel/bar, and carousel/text. However, for the x-comparison tasks (Task 4), the better combinations of RSVP and TYPE are shelf/text, collage/text, keyhole/bar, shelf/bar, and the worse combinations are convention/table and collage/table. Interaction plots of mean response time for RSVP×TYPE, RSVP×TASK, and TYPE×TASK are shown in Figure 9. The results of Experiment 2 showed that response times for the finding of maximum values (Task 3) were significantly faster than for graphs (Figure 9), and there was no significant difference between texts and tables. It is also consistent with the findings of Meyer, et al. [3]. However, for the x-comparison tasks (Task 4), response times were faster for texts and bars than for line graphs and tables (Figure 9) which reveal different findings with those of Meyer, et al. [3]. This is also inconsistent with the results that "Tables had an initial advantage over graphs for all tasks, and there were complex interactions between the variables."[2].

4 Conclusions

In summary, the results of experiments clearly indicate that the performance of various data extraction tasks depends on the RSVP mode and type of display used. The 2-way interaction of RSVP×TASK and TYPE×TASK will influence the visual decoding performance for the reading of exact values and trends in Experiments 1. The better designs of RSVP are convention, keyhole, and shelf modes for the tasks of finding exact values and keyhole, carousel, and convention modes for the tasks of identifying trends. Moreover, the 3-way interaction of RSVP×TYPE×TASK will influence the visual decoding performance for the tasks of x-comparison and maximum in Experiment 2. The better designs of combinations of RSVP and TYPE are convention/text, shelf/table, keyhole/text, and collage/table for the tasks of finding maximum values and shelf/text, collage/text, keyhole/bar, shelf/bar for the tasks of comparing values.

Acknowledgements. The support of the National Science Council, Taiwan (grant NSC 94-2213-E-029-014) is gratefully acknowledged.

References

1. De Bruijn, O., Spence, R.: Patterns of Eye Gaze during Rapid Serial Visual Presentation. In: Proceedings of Advanced Visual Interfaces, Trento, Italy (May 2002) 209–217 (2002)
2. Meyer, J.: Performance with Tables and Graphs: Effects of Training and a Visual Search Model. Ergonomics 43(11), 1840–1865 (2000)
3. Meyer, J., Shinar, D., Leiser, D.: Multiple Factors that Determine Performance with Tables and Graphs. Human Factors 39(2), 268–286 (1997)

4. Tse, T., Marchionini, G., Ding, W., Slaughter, L., Komlodi, A.: Dynamic key frame presentation techniques for augmented video browsing. ACM. In: Proceedings of Advanced Visual Interfaces, pp. 185–194 (1998)
5. Wittenburg, K., Forlines, C., Lanning, T., Esenther, A., Harada, S., Miyachi, T.: Rapid Serial Visual Presentation Techniques for Consumer Digital Video Devices. In: Proceedings of HCI International 2003 (June, Crete, Greece) Lawrence Erlbaum, vol. 5, 115–124 (2003)
6. Box, G.E.P., Cox, D.R.: An Analysis of Transformations. Journal of the Royal Statistical Society B 26, 211–243 (1964)

A Large Scale Study of English-Chinese Online Dictionary Search Behavior

Yong Liu and Jianmiao Fan

Indiana University, Bloomington, Indiana, USA
{yonliu,jfan}@dict.cn

Abstract. This paper presents a large scale study of user search behavior on a popular online dictionary website. Our goal is to understand the current status of online dictionary, especially English-Chinese dictionary user search behavior by analyzing 10 million queries on Dict.cn website during the period form late 2006 to early 2007. We believe our findings will help traditional dictionary publishers and online dictionary providers to improve their existing products and services, and further develop innovative services to better serve the unique needs of their users.

Keywords: Online dictionary, search interface, English-Chinese dictionary.

1 Introduction

Hundreds of millions of people in China are learning and using English on a daily basis. Many of them happen to be Internet users. The huge demand on English learning and the well established Internet infrastructure make online English-Chinese dictionary services more and more popular today. The goal of this study is to better understand the search behavior of online dictionary users. The study results can be used to help paper dictionary publishers and online dictionary providers to improve the existing products and services, and hence meet the unique needs of their users.

Dict.cn is a leading English-Chinese online dictionary website offering the largest English-Chinese dictionary database (over 3 million terms) to Internet users for free. It ranks among top 3000 most popular websites globally and serves over 1.5 million queries worldwide daily. In this study, we present analysis on Dict.cn's search logs collected through its Web interface, Web API, MSN bot, and WAP interface. The search log used in this study consists of 10 million queries randomly sampled during a three month period in late 2006 and early 2007. Both English and Chinese queries are included in this data set. Each search log entry contains fields including the user No., the query, the corresponding translation and a timestamp. All of our data is strictly anonymous, which means we maintain no data that can match a user with an identity. All of the results we report in this paper are aggregate statistics.

This study covers basic query statistics such as query length, queries per day, and repetitive queries. More sophisticated analysis is also conducted to examine the relation between users and their search patterns. For instance, one question this study tries to answer is how users forget what they learnt from a dictionary in a certain

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 389–398, 2007.
© Springer-Verlag Berlin Heidelberg 2007

period. Another important question is that if it is possible to classify users according to their search behavior. By answering these research questions, we believe this study will help paper dictionary publishers and online dictionary providers to achieve a better understanding of their customers' search behavior and hence improve their existing products or services.

The rest of this paper is organized as follows. Section 2 discusses some related work in user search behavior analysis field. Section 3 introduces multiple search interfaces provided by Dict.cn. Section 4 presents our dictionary query analysis based on a large scale search log data set. Finally section 5 gives a conclusion.

2 Related Work

A lot of research has been done on user search behavior analysis on general purpose search engines [2, 3, 6, 7]. These research efforts extracted user search patterns by studying a large set of search logs. The analysis results can be later used to improve the service quality of search engines. However, only a limited number of literatures discussed online dictionary user search behavior in the past decade.

In [8], Bergenholtz, H. et al. discussed using search logs to analyze user behavior in using an online Danish dictionary and how the results could be used to improved online dictionaries. In [9], De Schryver, et al. showed how a combination of search log analysis and the processing of formal online feedback forms may lead to improving dictionary contents. The same authors also studied the relation between search frequency and corpus frequency of online dictionary lemmas, and claimed their findings can be used to guide the compilation of future dictionaries [10]. However, most of these previous research efforts were based on a limited scale of online dictionary search logs involving only several thousand website visitors and tens of thousands queries. Such limitation made the reported statistical analysis results less convincing. Another thing the previous research failed to examine is the similar behavior pattern among multiple individual online dictionary users. Compared to traditional paper dictionaries, online dictionaries have great potential to facilitate the interactions among users during their searching and learning process. Study on the relations among individual users may help online dictionary providers to develop innovative services to better meet the unique needs of their users.

Different from previous research efforts on online dictionary search log analysis, our study is based on a large scale search log data set which came from a popular online dictionary website extensively used in real world. We focused not only on query statistics analysis, but also on user clustering by similar search behavior.

3 Dict.cn Search Interfaces

Dict.cn features four major dictionary search interfaces, which are Web interface, Web API, MSN bot, and WAP interface. The Web interface (a in Fig. 1) is Dict.cn's most visited search interface which Internet users access directly through a Web browser. Dict.cn also provides a MSN bot (b in Fig. 1), which users can add to their MSN messengers as a friend, to enable dictionary service via a chat interface. Dict.cn's WAP interface (c in Fig. 1) serves queries sent from cell phones while the

users are on the go. Last but not least, the Web API (d in Fig. 1) provided to third party developers by Dict.cn allows users to access Dict.cn's dictionary service via desktop software, toolbars, and browser plugins.

In the entire search log used in this study, 88.1% of queries came in through the Web interface. Besides, 8.7% of queries were received from the Web API and 2.6% from the MSN bot. The WAP interface contributed least to Dict.cn, which only received 0.4% of the total queries.

Fig. 1. Dict.cn Query Interfaces

4 Query Analysis

In this study we examined characteristics of user search behavior on Dict.cn. The six aspects we covered were query length, query difficulty, query distribution, queries per day, repetitive queries, and query based social network analysis.

4.1 Query Length

The entire search log consists of two types of queries – the English queries which were submitted by users looking for the corresponding Chinese interpretations and the Chinese queries, vice versa. For the English queries, we found the average number of words per query to be 1.67 (median = 1, max = 61, standard deviation = 1.4), with an average number of 10.87 (median = 9, max = 100, standard deviation = 7.79)

characters per query[1]. For the Chinese queries, the average number of Chinese characters per query was 5.39 (median = 4, max = 50, standard deviation = 4.92) [2]. Figure 2 shows the length distributions of both English and Chinese queries.

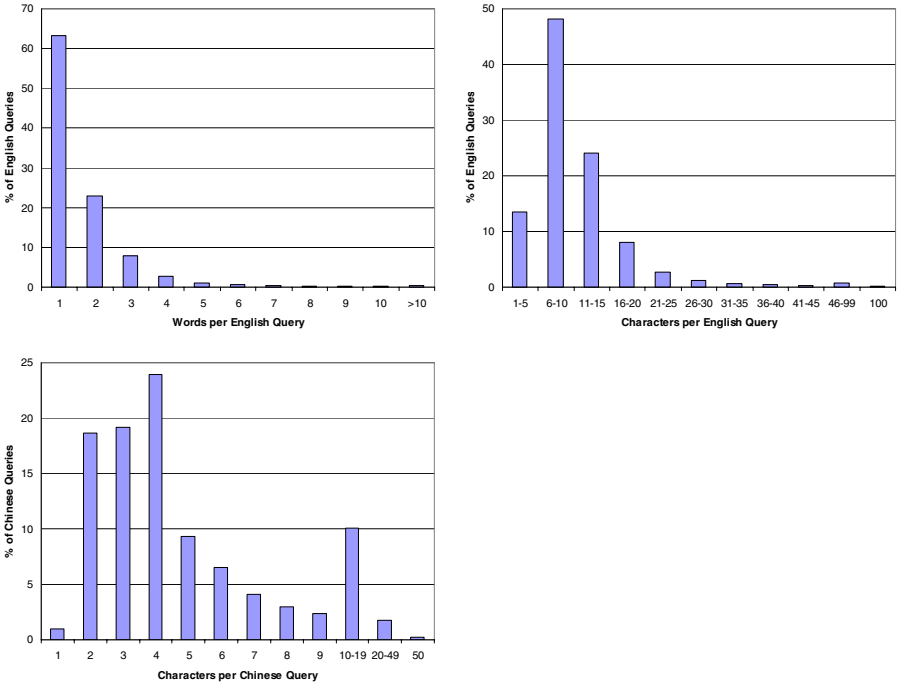


Fig. 2. Distribution of number of words and characters per English query and number of characters per Chinese query for Dict.cn queries received from all interfaces

It is interesting to compare the average query length between Dict.cn and general purpose search engines. Some published statistics suggested the average number of words per English query for search engines to be 2.35 [2] and 2.6 [3]. It is reasonable that the dictionary queries are much shorter on average since different from a search engine, words, not phrases, are usually the building blocks of an online dictionary.

Four-character query is the most popular Chinese queries according to Figure 2. However, previously reported average word lengths in various Chinese corpora are all less than 3 characters [4, 5]. This discrepancy suggests that online dictionary users tend to use combinations of two or more Chinese words to form search queries.

4.2 Query Difficulty

Previous research suggested that frequency of use is a valid measure of English word difficulty and can be utilized in vocabulary test construction and translation [1]. In

[1] The maximum number of characters in an English query allowed by Dict.cn is 100.
[2] The maximum number of characters in a Chinese query allowed by Dict.cn is 50.

this study we used this metric to examine the difficulty distribution of English queries on Dict.cn. The query difficulty measure *D(x)* used in this study is defined as following. As can be seen, the smaller *D(x)* is, the harder query *x* is.

$$D(x) = \text{sqrt}(\text{times of query } x \text{ submitted/number of total queries}). \quad (1)$$

In the search log used in this study, 60.3% of the English queries were submitted by users located in China, and 39.7% were from users outside of China. The average query difficulty of China based English queries is 8.22×10^{-3}(median = 0.0081, max = 0.00246, standard deviation = 0.0028). And the average query difficulty of non-China based English queries is 7.95×10^{-3} (median = 0.0076, max = 0.0246, standard deviation = 0.003). This means that users outside of China tend to submit more difficult English queries. This result is not surprising considering the fact that these users usually have better English language capability than Chinese domestic users.

Figure 3 shows the distribution of averaged English query difficulty for both China and non-China based user groups. This again demonstrates that Chinese domestic users tend to submit easier English queries to Dict.cn. For the better understanding of the meaning of query difficulty, here are some examples on query difficulty: *D*("imprimatur") is 4.5×10^{-3}, *D*("privacy") is 8.3×10^{-3}, and *D*("next") is 1.8×10^{-2}.

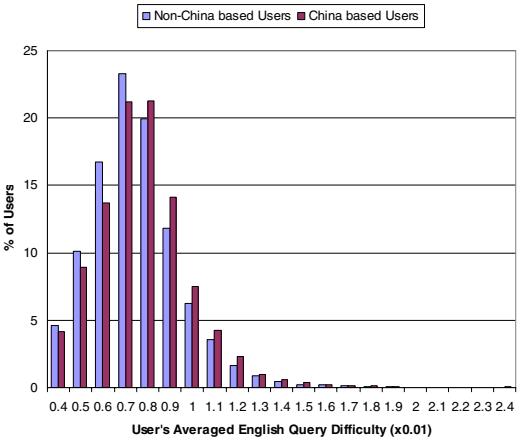


Fig. 3. Difficulty level of English queries by users

4.3 Query Distribution

More than one million queries are submitted to Dict.cn daily. However, the number of unique queries submitted is much smaller. To study the variation in the queries, we examined the percentage of the total query volume that was covered by top-N unique queries. We picked the mostly used 1000 English and Chinese queries from the dictionary search log and calculated the distribution of top 1 to 1000 queries.

As Figure 4 shows, English queries vary more than Chinese queries do. The percentage of total searches covered by the top 100 Chinese queries is 4.3% while the percentage covered by the top 100 English queries is only 2.4%. However, the difference between the coverage of top 1000 English and Chinese queries is not that much (13.1% vs. 15.9%). This suggests that Chinese queries focus more heavily on the top couple of hundred unique queries.

Previously a similar study reported that top 1000 search engine queries accounted for approximately 6% of all queries [6]. The same study also reported that top 1000 mobile search queries accounted for around 22% of all mobile queries. These findings suggest that general purpose search engine queries are much more diversified than dictionary queries, whereas mobile search queries are less diversified. This suggests that the usable contents on the existing mobile Internet is still not as rich as the contents provided by a large scale online dictionary.

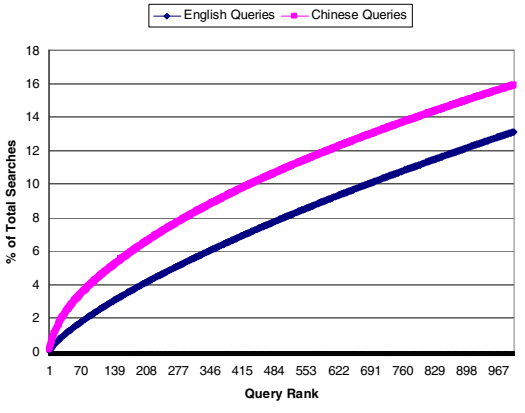


Fig. 4. Cumulative percentage of total searches accounted for by the top 1000 English and Chinese queries

4.4 Queries Per Day

As part of our study, we also examined the distribution of the number of queries submitted by users per day. Figure 5 illustrates the study results for both China based user groups and non-China based user groups. As can be seen, the largest user portions in both groups are those who submitted only one query per day. However, both groups feature a long tail in the figure. There are a fairly large portion of users who submitted more than 20 queries per day in both groups. The more queries a user submits, the better his or her search behavior can be profiled. Based on a user's profile, more customized services can be provided in the future.

Figure 5 also shows that non-China based users tended to submit more queries per day. The reason might be that these users had stronger needs to use a dictionary since they stayed in a non-Chinese language environment.

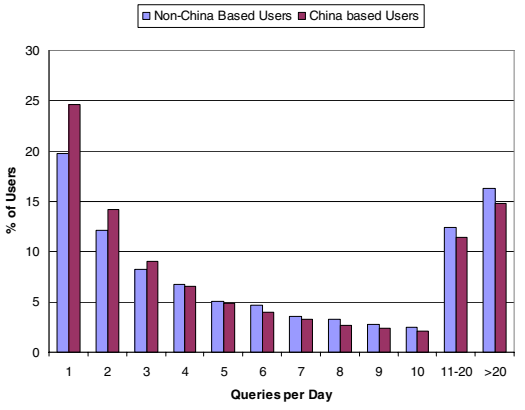


Fig. 5. Queries per Day by China and Non-China based Users

4.5 Repetitive Queries

People forget things. And this happens a lot during the learning process. To study how fast our users forget their newly learnt words, we examined the number of repetitive queries a user launched in a 10 day period. Figure 6 shows the result.

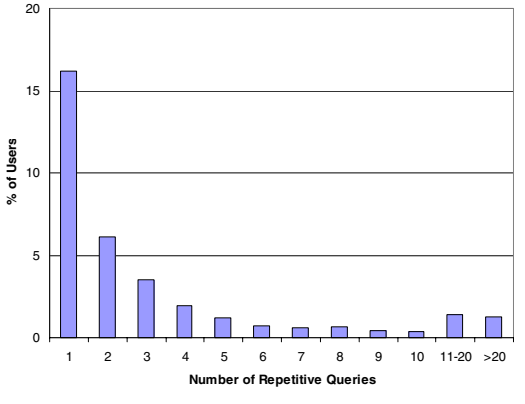


Fig. 6. Number of repetitive queries by one user in 10 days

We found that 34% of users at least searched for an identical word or phrase twice in 10 days, and 2.7% of users searched for more than 10 identical words. This finding explains why dictionaries always sell, and why a dictionary website can survive without any updates for a long period.

4.6 Query Based User Clustering

An identical query may be submitted to an online dictionary by different users during a certain period. The users who did this may share the same background or the same

interests. For instance, users who searched for the phrase "hyper threading" have a high possibility of either having computer science background or being interested in computer technology. Also, users who submitted the same queries may have similar language capability. If we can find those similar users (in terms of background, interest, language capability, etc.) based on the search logs, we will be able to provide innovative services to online dictionary users such as collaborative study.

In this study we randomly picked 1022 users who submitted at least 30 queries in a one month period. We examined the number of users who shared a certain number of queries with at least one other user. As can be seen from Figure 7, more than 20% of users shared at least 10 queries with one or more other users. Around 10% users shared at least 15 queries. This result demonstrates the feasibility of finding a similar peer for at least 10% of total users since 15 identical queries can largely guarantee that two users have some similarity between them. Given a longer period of the search log collection, the matching of two similar users can be more accurate.

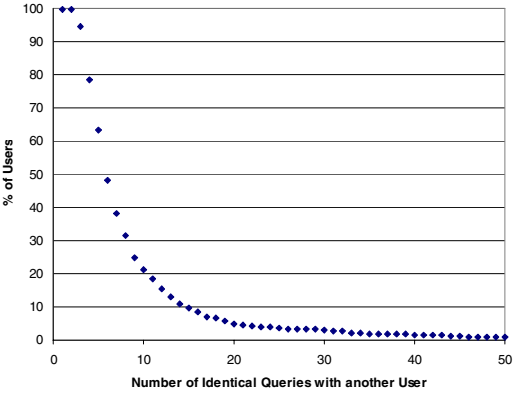


Fig. 7. Number of identical queries with at least one other user

We also examined the possibility of clustering users into separate user groups to maximize the similarity within each group. By applying a fast clustering algorithm published in [11], we clustered the before mentioned 1022 users into 125 groups. Figure 8 (a) shows a query similarity matrix of users where each spot denotes the corresponding pair of users who had at least 3 identical queries. Since the users are randomly arranged along two axes, the similarity matrix looks totally random. Figure 8 (b) shows the 3D similarity matrix in a diagonal perspective. As can be seen, the majority of the similarity pairs have 3 or 4 identical queries. Figure 8 (c) shows what the similarity matrix is like after users are re-arranged along two axes according to the 125 clustered groups (users in the same group are placed adjacent to each other on x and y axis). We can see now most points are gathered along the principal diagonal of the similarity matrix. Figure 8 (d) provides a better view where we clearly see the similarity points, especially those with 5 or more identical queries, are now gathered tightly along the diagonal. This demonstrates that the clustering algorithm works very well on the online dictionary users based on their queries.

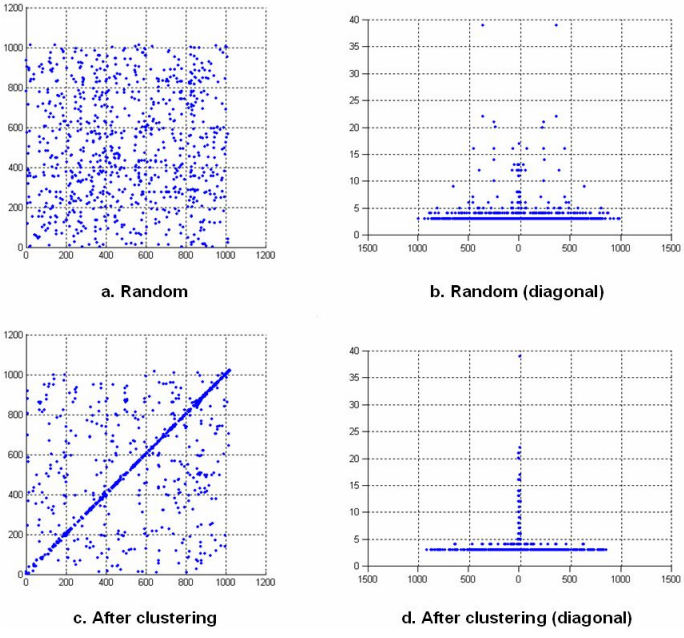


Fig. 8. Query based user clustering

5 Conclusions

In this paper we examined online dictionary user search behavior based on a large scale search log data set of a leading online English-Chinese dictionary website. The study results suggest that:

- On average, online dictionary queries are shorter than search engine queries.
- English queries submitted by Chinese domestic dictionary users are less difficult than those submitted by non-China based users.
- Online dictionary queries are less diversified than search engine queries.
- Non-China based users query more per day on average.
- A large number of users submit repetitive queries in a short time period.
- Online dictionary users can be well clustered into separate groups according to the similarity of their search behavior.

Although we have presented some findings on dictionary query based user clustering in this paper, we believe there are still many interesting open research issues in this topic. This will be the subject of our next large scale study on online dictionary user search behavior.

Acknowledgments. Here we would like to gratefully acknowledge the help of all Dict.cn users. This paper would not be possible without their input.

References

1. Tamayo, J.M.: Frequency of Use as a Measure of Word Difficulty in Bilingual Vocabulary Test Construction and Translation. Educational and Psychological Measurement 47(4), 893–902 (1987)
2. Silverstein, C., Henzinger, M., Marais, H., Moricz, M.: Analysis of a Very Large Web Search Engine Query Log. SIGIR Forum 33(1), 6–12 (1999)
3. Spink, A., Jansen, B., Wolfram, D., Saracevic, T.: From E-Sex to E-Commerce: Web Search Changes. IEEE Computer 35(3), 107–109 (2002)
4. Nie, J.Y., Gao, J., Zhang, J., Zhou, M.: On the Use of Words and N-grams for Chinese Information Retrieval. In: Proceedings of the Fifth International Workshop on Information Retrieval with Asian Languages, Hong Kong, China, pp. 141–148 (2000)
5. Xu, J., Zens, R., Ney, H.: Do We Need Chinese Word Segmentation for Statistical MachineTranslation? In: Proceedings of the Third SIGHAN Workshop on Chinese Language Learning,Barcelona, Spain, pp. 122–128 (2004)
6. Kamvar, M., Baluja, S.: A Large Scale Study of Wireless Search Behavior: Google Mobile Search. In: Proceedings of the SIGCHI conference on Human Factors in computing systems (2006)
7. Wen, J.R., Nie, J.Y., Zhang, H.J.: Clustering User Queries of a Search Engine. In: Proceedings of International WWW Conference (10). Hong-Kong (2001)
8. Bergenholtz, H., Johnsen, M.: Log Files as a Tool for Improving Internet Dictionaries. Journal of Linguistics 34, 117–141 (2005)
9. De Schryver, G.M., Joffe, D.: On How Electronic Dictionaries are Really Used. In: Proceedings of the Eleventh EURALEX Interna-tional Congress, Lorient, France, pp. 187–196 (2004)
10. De Schryver, G.M., Joffe, D., Joffe, P., Hillewaert, S.: Do Dictionary Users Really Look Up Frequent Words? – On the Overestimation of the Value of Corpus-based Lexicography. Lexikos 16, AFRILEX-reeks/series 16, 67–83 (2006)
11. Frey, B., Dueck, D.: Clustering by Passing Messages between Data Points. Science DOI: 10.1126/science.1136800 (2006)

Design of Computer Integrated Safety and Health Management System

Hunszu Liu

Industrial Engineering and Management, Ming Hsin University of Science and Technology,
No. 1 Hsin-Hsing Road, Hsin-Fong, Hsinchu, Taiwan
hliu@must.edu.tw

Abstract. Diverse safety and health operation data collected and stored in different departments have not been fully integrated and utilized by managers due to poor design of safety and health information system. The safety and health management system can be depicted, conceptually, as an organic system with circulation of information flow which carries required data and information to specified workers and initiates appropriate responses respectively. This study is targeting to solve the problems of current safety and health management system through the integration of human information processing theory, certified safety and health management assessment guidelines and regulations, and IT techniques. The objective of this study is to propose a framework of computer integrated safety and health management system.

Keywords: Safety and health management, safety information, computer integrated management system.

1 Introduction

Various safety and health management system were developed to meet the requirements of government law and regulations. Among them, two types are most commonly used, the traditional method of safety, and the methods and philosophies of quality in conjunction with safety. The traditional methods of safety established standard procedures to direct and control workers. Management programs are in place to modify the behavior of workers, to motivate their participation, and to reward their compliance. This approach centers exclusively on the technical requirements and on obtaining short-term results. Another shortcoming of this approach is its isolation with the rest of functions of organizations [1], [2]. In recent year, the concept of integrate safety and quality is widely accept by safety practitioners to reinforce the weakness of traditional approach. Total safety management (TSM), written and practiced using the principles of total quality management (TQM), was introduced to industry [3], [4], [5], [6]. However, these management programs do not always improve the results of safety. This is because diverse safety and health operation data collected and stored in different departments have not been fully integrated and utilized by managers due to poor design of safety and health information system.

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 399–405, 2007.
© Springer-Verlag Berlin Heidelberg 2007

Decisions are made based on insufficient information, thus decreasing the robustness of system safety barriers.

The safety and health management system should not only meet the regulation requirements, which are the minimum requirement, but also act as an effective information processing system which means it can handle and store stimuli generated from system components and capable of performing perception, decision and response selection, and response execution functions. In other words, the system should function as an organic system with circulation of information flow which carries required data and information to specified workers and initiates appropriate responses respectively. This study is targeting to solve the problems of current safety and health management system through the integration of human information processing theory, certified safety and health management assessment guidelines and regulations, and IT techniques. The objective of this study is to propose a framework of computer integrated safety and health management system.

2 Background Information

The advent of technology has significantly increased the scope and paces of safety and health management activities. Information, intelligence, and computer techniques were widely applied in industry environment. These techniques reduce the burden of manual operations such as data collection, environment detection, and data transformation. More activities can be included in the safety and health management program. But the amounts of data generated from these devices are also increased. As a result, the workloads of safety and health personnel have been changed from physical activities to psychological activities. Most manual works have been replaced by mental works. The success of this change relies heavily on the design of safety and health management system (SHMS).

A well designed safety and health management system should meet the organization needs. The magnitude of organization and manufacturing characteristics are two important factors to decide the scope and complexity of SHMS. While multination corporations can spend enough resources to establish and maintain complex SHMS, small enterprise barely has designated personnel in charge of safety and health issues. To balance the differences between organizations and set an objective criteria for assessing the SHMS, safety and health standards or guidelines, national or international, are developed. Among them, Process Safety Management (PSM), Voluntary Protection Program (VPP), Responsible Care (RC), International Safety Rating System (ISRS) [7], [8], [9], and Occupational Health and Safety Assessment Series 18000 (OHSAS 18000) [10], [11] are the most popular in local industry. These standards or guidelines focus on the establishment of safety and health management system mechanisms. Fulfilled the requirements and passed the assessment process indicate organizations are embedded with a certified safety and health management system. It is argued that these mechanisms, accompanied by scheduled audit process, will continually improve organization safety status.

One problem of these approaches is that these methods are developed and tested through pragmatic practical approaches, which have not been scientifically validated.

Another problem is that these requirements are either too general to be used directly by plant managers to construct its safety and health management system or

too strict to be used by small companies. The safety and health management system must fulfill the requirements of general accepted standards as well as industrial managers' expectations. Therefore, the safety and health practitioners' opinions must be incorporated into the developing process. The third problem is that these systems usually act as an isolate system. In order to fulfill the requirements of these systems in very short time, organizations usually form a temporary project team in charge of the implementation process. The project teams members come from different departments develop appropriate measurements through extra-working hours. Their performances are rated based on the final audit results. It is likely to ignore activities which are not required by those systems. Many safety and health activities current carried out within the organization will not be included in the system. As a result, these systems usually operated without integration with existing safety and health system.

3 Methodology

The process of developing computer integrated safety and health management system is divided into three stages. The objective of stage I is to develop the components of system. Stage II focuses on the program development. Stage III is the implementation and risk assessment stage. This paper describes works performed in stage I. In order to explore the components of a well designed safety and health management three most widely implemented safety and health management system in local industry, including Process Safety Management (PSM), Voluntary Protection Program (VPP), and Occupational Health and Safety Assessment Series 18000 (OHSAS 18000), are studied. The functional requirements of these components and corresponding activities are identified through interviewing with industry personnel. Questionnaires were developed through literature review. The components are classified into three categories, including Occupational Health and Safety Condition (OHSC), Occupational Health and Safety Management (OHSM), and Occupational Health and Safety Operation (OHSO). A Total of thirty-two items was identified (Table 1).

The Analytic Hierarchy Process (AHP), developed by Thomas L. Saaty, is used as the tool to analyze the collected questionnaires. Each participant was required to provide judgments about the relative importance of each component.

The questionnaire is divided into two parts. Part I. pair wise compares the priorities of the three categories: OHSC, OHSM, and OHSO. Part II. pair wise compares the priorities of the components within each category respectively. The comparison scale for the importance of criteria is classified into nine levels and converted into numerical level from 1-9: (from equally important to extremely more important).

4 Results and Discussions

Thirty senior safety and health experts with at least ten years of related experiences were asked to answer the questionnaires. Eighteen questionnaires were collected and analyzed to identify the necessity and importance of these criteria. The analyzed results indicate that the priorities for three safety and health management system categories are: A. (OHSC, 1.36), B. (OHSM, 1.68), and C. (OHSO, 0.96). The OHSC

reflects the general condition of the whole safety and health management system and the OHSO indicates the operational level status. The results indicate that for plant managers the image of safety and health management system and operational status are their major concerns.

The rankings of safety and health management system components are listed in table 2. The implementation of safety and health program and employee safety behavior are the most important component for a well designed S&H management

Table 1. Safety and Health Management System Component

Categories	Components
A.OHSC	A1:Occupational injury and disease condition
	A2:Operational environment quality
	A3:Working Plant Health and Safety Condition
	A4:Legal Compliance
	A5:Employee Health Inspection Condition
	A6:Employee Safety Awareness
	A7:Employee Safety Behavior
	A8:Mechanical Equipment Safety Status
	A9:Safety Climate
B.OHSM	B1:Management Review
	B2:Safety Audit
	B3:Achieving Objectives
	B4:Health and Safety Improvement Program
	B5:Employee Suggestion Program
	B6:Health and Safety Promotion Activities
	B7:Corrective and Preventive Action
	B8:Customer or Community Complaints
C.OHSO	C1:Top Management Inspection
	C2:Implementation of the Training Program
	C3:Implementation of Health and Safety program
	C4:Implementation of the Self Inspection Program
	C5:Implementation of the Emergency Response Program
	C6:Implementation of the Occupational Injury and Disease Prevention Program
	C7:Dangerous Machinery and Equipment Operation Condition
	C8:Occupational Injury and Disease Report and Treatment Condition
	C9:Operational Control Condition
	C10:Contractor control Condition
	C11:Consultation and Communication Condition
	C12:Personal Protective Equipment Operating Condition
	C13:Toxic and Danger Substances Management Condition
	C14:Implementation of the Process Safety Assessment and Risk Assessment
	C15:Measurement Equipment Calibration Condition

Table 2. Ranking of safety and health management system performance indicators

rank	performance indicators	priority	rank	performance indicators	priority
1	C3	1.82	17	A3	0.94
2	A7	1.78	18	A5	0.92
3	C5	1.48	19	A1	0.9
4	B3	1.41	20	B7	0.8
5	C1	1.36	21	C12	0.8
6	A6	1.33	22	C13	0.8
7	A4	1.21	23	C9	0.79
8	B2	1.2	24	A2	0.76
9	C14	1.15	25	B1	0.74
10	C6	1.13	26	C11	0.66
11	C7	1.11	27	A8	0.62
12	C2	1.08	28	A9	0.54
13	C4	1.01	29	C8	0.48
14	B5	1	30	B8	0.45
15	C10	1	31	B6	0.41
16	B4	0.99	32	C15	0.33

system. The measurement equipment calibration condition and health and safety promotion activities are the least important component. These results concur with previous conclusion that the image of safety and health management system and operational status are their major concerns.

5 Discussions

As the manufacturing system progressing toward the lean manufacturing system era, the tolerance for deviation from normal operations becomes more tightly than ever before. The unexpected breakdown of safety and health management system can initiate chain reactions of business interruptions, once the system safety barriers are unable to protect the disturbances from either technical or human failure, which might endanger business survivals for companies involved. In order to ensure the smooth operation of manufacturing activities, safe working environment must be achieved through robust design on defense in depth and effective safety management system. However, the working environment will not stay the same. The changing manufacturing scenario created by aging facilities and inconstant human performance has put the effectiveness and efficiency of deployed program in doubt. As a result, to ensure the proper implementation of safety and health management system and maintain the on-going process of fulfilling the organizational needs, it is necessary for companies to have a well designed safety and health management system which can serve managers' needs.

Traditionally, company managers receive safety and health related information through the feedback of the company's safety and health accident records. As a result, companies can only realize the safety and health problems when something went wrong. To reduce the consequence of the trial-and-error approach adopted currently by many organizations, it is necessary to develop a computer integrated safety and

health management system, which can help company managers assessing the management system status.

The purpose of this study is to develop a framework of computer integrated safety and health management system. Related researches are reviewed and expert knowledge is incorporated. The study results can provide the system developer the capability to construct the safety and health management system.

6 Future Works

This study discusses partial research results of three years project. The components of a well designed safety and health management system are investigated. The next step is to analyze these activities and construct the information system through Business System Planning (BSP) technique. Components are grouped together based on the expected functions performed under the human information processing model. The data collected at each department or working site through sensors or manual entry play as stimuli, which represent current system status of the information processing system. Components are designed to perceive and process these stimuli. Functions of tools to help decision maker select proper responses are identified. IDEFØ technique is adopted to model the decisions, actions, and activities of safety and health management system. In order to maintain the effectiveness of safety and health management the Plan-Do-Check-Action (PDCA) principle are embedded in the system. This means labor intensive works are needed to process system operation information. Computer integrated safety and health management system can not only alleviate the burden of paper works but also provide efficient way of accessing these data.

Acknowledgments. The authors would like to thank the National Science Council for supporting this project (project # NSC#95-2221-E-159-005)

References

1. Smith, T.A: Will safety be ready for workplace 2000. Professional Safety 41(2), 37–38 (1996)
2. Herrero, S.G., Saldan, M., et al.: From the traditional concept of safety management to safety integrated with quality. Journal of Safety Research 33, 1–20 (2002)
3. Rahimi, M.: merging strategic safety, health and environment into total quality management. International Journal of Industrial Ergonomics 16, 83–94 (1995)
4. Weinstein, M.H.: Improving safety program through total quality, Occupational Hazards, 58(8) (1996)
5. Goetsch, D.: Implementing Total Safety Management. Prentice Hall, New Jersey (1998)
6. Sznaider, B.A.: TQM can make plants safer. Manufacturing Engineering 121(5), 144 (1998)
7. Bird, F.E., Germain, G.L.: Practical Loss Control Leadership, International Loss Control Institute, Georgia, USA (1990)
8. Eisner, H.S., Leger, J.P.: The International Safety Rating System in South African Mining. Journal of Occupational Accidents 10, 141–160 (1988)

9. Gaunt, L.D.: The effect of the International Safety Rating System on organizational performance. Centre for Risk Management & Insurance Research. Georgia State University, Research Reports Service. 89-2. Atlanta, Georgia (1989)
10. British Standards Institution,BS 8800, Guide to Occupational Health and Safety Management System (1996)
11. BS 8800: Guide to Occupational health and safety management systems,BSI, UK (1996)

Sensor-Driven Adaptation of Web Document Presentation

Spyros Nathanail, Vassileios Tsetsos, and Stathes Hadjiefthymiades

Pervasive Computing Research Group, Dept of Informatics and Telecommunications
University of Athens, Panepistimiopolis, Ilissia, 15784, Greece
{s.nathanahl,b.tsetsos,shadj}@di.uoa.gr

Abstract. The increasing variety of user device technologies has raised the necessity for ubiquitous content provision, which is characterized by "intelligent" content delivery to end users, irrespectively of their device. Moreover, changes in the user environment create the necessity of appropriately adapting the content and its presentation so that it can be more easily perceived. In this paper, we present a multimodal and adaptive system for Web user interfaces that deals with these challenges. Using as input a generic and abstract form of Web content, our system transforms it to a specific data format suitable for the client device. In addition, it uses contextual information, gathered from the user environment through a wireless sensor network, in order to present the served content in an optimal way under the current environmental conditions.

Keywords: adaptive user interfaces, abstract user interfaces, context-awareness, wireless sensor networks.

1 Introduction

According to the Ubiquitous Computing paradigm, which is increasingly penetrating our everyday lives, information processing and communication tasks should be available to roaming users with limited effort from their part. Central concepts in this paradigm are, among others, seamless support for device heterogeneity, advanced user interfaces and context-awareness. Hence, the ability of adapting the information to many different forms, according to the user's general context, is deemed crucial for smart environments and services. Similarly to the need for information adaptation there also arises the need for *context-aware presentation* of the delivered information to the users. The term "context" refers to everything that affects the interaction between a user and her computing environment, such as device capabilities, user status/profile and environmental conditions [8].

The vision of context-aware content presentation is what has mainly motivated our work. In this paper, we describe the design, implementation and performance evaluation of a system, named Chameleon, that performs such adaptation. An indicative use case for the proposed system is the following:

An emergency response team using various hand-held devices (e.g. Tablet PCs, PDAs) needs information that can be updated and adapted rapidly and accurately,

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 406–415, 2007.
© Springer-Verlag Berlin Heidelberg 2007

according to their surrounding environment. Since that team usually operates under stress and extreme environmental/weather conditions, changes in factors such as light and noise levels must be treated by the application, without extra effort and attention from the team members. Hence, the proposed system a) transforms the content to a format appropriate for each member's device, and, b) adapts the content presentation elements (colour, size, etc.) to the environmental conditions, provided that the adaptation rules and the sensing infrastructure have been set.

Obviously, Chameleon can also facilitate the Web content provisioning process for people facing various disabilities (e.g., elderly people or people with visual impairments).

In general, an adaptive and multimodal user interface system that adheres to the Ubiquitous Computing paradigm should address three requirements:

a) access to the same content from devices using different platforms and having different capabilities,
b) adaptive formatting of the content in order to address changing contextual conditions,
c) seamless provision of the above functionality (i.e., transparently to the end user)

The proposed system fulfills these requirements by extending the functionality of a web server so as to deliver content in a format tailored to the user device capabilities and the contextual status of the mobile or stationary users. Specifically, Chameleon senses the environmental conditions near users through wireless sensor networks, and properly adjusts the formatting/presentation of the requested content, according to predefined presentation rules. Such rules may be the same for all users or adjusted appropriately on a per-user basis.

The implementation of such system would be of little value and utility without the recent advances and wide acceptance of Wireless Sensor Networks (WSN). Such networks enable remote sensing and can be attached to moving objects, such as handheld devices or wearable computers. Hence, they provide a ubiquitous sensing infrastructure that can be transparently incorporated in pervasive and mobile services.

The organization of the paper is as follows. In Section 2, we discuss the overall architecture and functionality of the system as well as describe its end-to-end workflow. In Section 3, we further discuss some selected implementation issues. In Section 4, the setup and the results of a performance evaluation are presented. Finally, in Section 5, we present some related work and the paper concludes with some directions for further research.

2 Chameleon Architecture and Functionality

Chameleon consists of three modules. The *M-module* provides the multimodality functionality. The content is initially described in an *abstract user interface language*, which only describes its structure and semantics. The M-module transforms such content according to a user interface language that can be supported by the user device (hereinafter referred to as *target language*). The transformed content contains also presentation-specific annotations. The *A-module* provides the adaptation

functionality of Chameleon. It is responsible for formatting the content resulting from M-module according to i) certain presentation rules, and, ii) the data received from the sensing infrastructure (i.e., WSN). The S-module is responsible for collecting and handling the sensor readings. These modules, along with the basic data flows, are depicted in Figures 1 and 2 and described in the following paragraphs.

A session begins when the user accesses the server from a web client (i.e., browser) and issues a certain request for content. For the rest of this paper, we assume that such content is a Web page. After the server has received the request, it retrieves the page in its original, abstract form and forwards it to the M-module.

2.1 System Modules

The M-module is responsible for the translation of the original, abstract web page into a final web page that suits the browser and device capabilities of the requesting user. The page transformation is performed by applying a set of XSL Templates [16] on the original page. The outcome is a page formatted in a target language (e.g., HTML) that will eventually be sent to the user device. The whole process is transparent to the user.

It should be noted that Chameleon does not send an abstract XForms document to the user and, thus, it does not burden the user's browser with the task of converting it to an appropriate target language. Such an approach would not permit the context-driven adaptation of the final page, since most end devices cannot retrieve and process contextual information locally. In the current implementation we support only HTML as a target language. Hence, we will not deal with multimodality issues in the rest of the paper, but focus on the sensor-driven presentation adaptation instead.

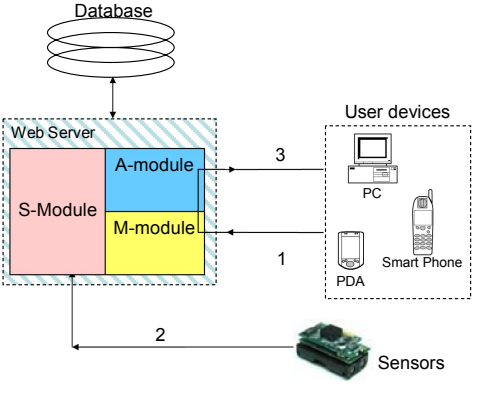


Fig. 1. System architecture and main data flows (1: User request, 2: Sensor readings, 3: System response)

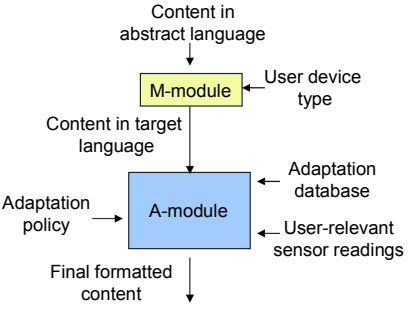


Fig. 2. Basic system functionality

After the M-module has performed the page transformation, the system forwards the processed page to the A-module. The latter is responsible for the adaptation and formatting of the page content based on the contextual information sensed from the user environment. Such information is retrieved through a wireless sensor network,

the sensors of which may be physically located on the user's device (hereinafter called *device sensors*), or can be deployed at the broader area of the user (hereinafter called *area sensors*). The sensors constantly communicate with the system through a middleware in order to report their readings (see also Section 3.1).Prior to deciding how the adaptation of the content will be performed, the system must select which sensors will affect the presentation adaptation. This is the responsibility of the S-module.

Upon system deployment, a permanent ID is assigned to each user device (e.g., PC, PDA, mobile handset). This ID relates the device to a) its device sensors and b) the location in which it resides. For the location estimation of a device, several positioning methods can be used, either for indoor or outdoor environments [7]. By knowing where the device is located, the system can determine which area sensors are relevant to each user. Hence, only that particular subset of device and area sensors (hereinafter called *user-relevant sensors*) are monitored, and their readings are stored to a database table accessible by the A-module.

2.2 Adaptation Through Abstract Policies

The rules that govern the reaction of A-module to sensor value changes are specified as Adaptation Policies (AP) expressed in an Adaptation Policy Language (APL), which we have defined. In particular, the system, during the adaptation of the requested page, checks the values of all user-relevant sensors against the rules specified through the AP. An AP is an XML document that assigns *groups of actions* to sensor ranges, with each action group containing one or more generic formatting commands (i.e., *actions*). Such generic commands are not bound to any target language, but are rather "rough instructions" for the final page layout. The AP is specified with abstract formatting commands instead of “target language” tags in order to support multimodality.

Sensors of the same type are grouped together by Chameleon, and a single value (e.g., the median of the individual sensor readings) is computed for each sensor group. We have decided to give priority to device sensors over area sensors, since the former are, in general, closer to the user and its device, and, thus, their readings are expected to be closer to the actual conditions of the user-device interaction. In Table 1 some indicative rules are described, where generic actions (e.g., LowContrast) are assigned to the various sensor value ranges. The actions fall into *attribute* categories which group together actions that affect similar aspects of formatting (i.e., “Size”, “Contrast”, “Volume”). This classification of abstract actions in Table 1 primarily serves the rule author of the system in organizing the rules.

The actual adaptation is performed through the mapping of abstract actions to CSS tags (see also Section 3.3). First, we create a table containing all CSS tags for each abstract action along with their default values (see Table 2).

Subsequently, we create the mappings between the abstract actions and the CSS tags (see Table 3). These mappings also specify the target language to which each transformation is referred, and a *relative modifier* that will be applied to the CSS tag value. For example, in the case of the action *LargeSize* the modifier applied to the

Table 1. Sample Adaptation Policy rules

Attribute / Sensor	Sensor Value	Contrast	Volume	Size
Luminance Sensor	0-5	LowContrast	-	MidSize
	5-7	-	-	-
	8-10	HighContrast	Loud	-
Velocity Sensor	0-3	-	Soft	SmallSize
	4-7	-	-	MidSize
	7-10	HighContrast	Loud	LargeSize

Table 2. Table of default formatting values

End Language	CSS Tag	Default Value
HTML	font-size	9
HTML	font-weight	400
VoiceXML	voice-volume	5
HTML	background-color	0xFFFFFF

Table 3. Mappings between AP actions and CSS tags

End Language	AP Action	CSS Tag	Modifier
HTML	LargeSize	font-size	+4
HTML	LargeSize	font-weight	+100
HTML	SmallSize	font-size	-2
VoiceXML	Soft	voice-volume	-2
VoiceXML	Loud	voice-volume	+3

CSS tag "font-size" is "+4". We already know from Table 2 that the default value for the "font-size" tag is "9". Therefore, if the system executes the *LargeSize* action, it will set a new value of "13" to "font-size".

The modifiers, of course, have to be numbers so that they can be added/subtracted. However, there is a solution for non-numerical modifiers as well. The values of such modifiers can be represented as *enumerators*, assuming the relevant CSS tags have *ordered* values (e.g. small, medium, large). Hence, instead of manipulating tag values, we add/subtract enumerator values. The final value of the tag is the value of the enumerator that corresponds to the number resulting from the addition/subtraction of the relevant modifiers. For example, consider the CSS tag "border-width" which may take the values "thin", "medium" and "thick". Since the values can be ordered, we represent them with the numerical values "1", "2" and "3", respectively. If the current tag value is "thin" and the following hypothetical rule applies, then the "border-width" will be set to "medium":

If velocity > 0.3m/sec then "border-width" = "border-width" + 1

When transformation is complete, the system has produced a set of CSS tags that will be applied to the end page. These tags are integrated in the page as an internal CSS style (see Section 3.3). This process is repeated every time the value of at least one of the user-relevant sensors changes, thus, causing the system to "push" the new version of the page to the client. This automated update is performed through AJAX technology [18].

It should be noted a this point that the default values in Table 2 and the modifier values in Table 3 can be further personalized based on the user profile. This can be achieved by adding an extra element in the rule antecedents: the user class. The implementation of such personalization has not incorporated in the Chameleon prototype yet, but it is rather straightforward. All that is needed is a way to capture/extract the user profile (e.g., possible limitations regarding the perception of user interfaces) and to rewrite the adaptation rules by taking this profile into account.

Finally, similarly to every rule-/policy-based system, there may occur the situation that conflicting actions should be executed. In such case, two categories of conflicts can be identified:

1. *Rules that affect different presentation elements in conflicting ways.* For example, a rule that increases the text font size in a web page and another one that decreases the size of buttons. This type of conflict should be addressed by the AP author, who should create rules that affect all elements on a page and not just a subset of them.
2. *Rules that affect the same presentation element in conflicting ways.* In order to resolve such conflicts, the APs should be designed so that the rule actions do not result in contradictory formatting instructions. This last type of conflict is addressed by the system through the use of relative modifiers for the CCS tags, since the final value of the CSS tags is computed by adding/subtracting the values of all relevant modifiers.

3 Selected Implementation Issues

3.1 Wireless Sensor Network and Middleware

For the development of a system prototype, a limited WSN was used. The sensor nodes were the Crossbow Mica2 motes [9], carrying light, noise and temperature sensors (with only the first being exploited). Each mote was executing the TinyDB application (part of the TinyOS suite [20]), which enables the programmer to issue SQL-like requests against the sensor network. The readings of the sensors were periodically forwarded to the S-module, where the database table with the sensor values was updated accordingly.

3.2 Abstract User Interface Languages

Another key decision was to use XForms [15] as the abstract user interface language. The main reasons for adopting XForms are:

1. XForms is a W3C standard. Hence, it is well defined, widely supported and not likely to be abandoned in the near future.
2. XForms separates data from presentation. This, practically, means that the content author does not have to get involved with page presentation details.

A number of other languages were considered prior to adopting XForms. These are the User Interface Markup Language [4], the Extensible Interface Markup Language [19] and the Alternative Interface Access Protocol [17]. A number of reasons

discouraged us from using the three above languages. For example, UIML is well-developed but does not separate data from presentation as strictly as XForms does. Moreover, both XIML and UIML require an abstract user interface vocabulary (a set of mappings between abstract and concrete elements) to be included in the page in order to render it. AIAP, on the other hand, does not impose the aforementioned restrictions, but is still under development from INCITS (InterNational Committee for Information Technology Standards) and is not yet a standard, thus limiting its potential support and usability.

3.3 Adaptation Through Cascading Style Sheets

As already mentioned, the pages are formatted using Cascading Style Sheets (CSS). Apart from being a well-established standard, the recent developments in its specification offer versatility and interoperability with mobile devices (CSS Mobile Profile) [10] and voice-enabled devices (CSS3) [11]. CSS Mobile Profile offers a variety of formatting commands that are specific to mobile devices (e.g., cell phones). CSS3 includes a speech module, which enables auditory formatting of a rendered page. The CSS flavor used by our system is *internal CSS*. In internal CSS the formatting instructions are embedded in a HTML tag (<STYLE>) in the page header.

Regarding the CSS update procedure, changes in page formatting are performed on the client-side through the use of the AJAX [18] programming model. More specifically, the client, using a XMLHttpRequest object, periodically queries the system regarding sensor value changes. If such changes have occurred since the last check, the system responds with the new CSS stylesheet elements, which replace the old ones. This way, formatting is performed seamlessly (i.e., without discarding any user data).

4 Performance Evaluation

4.1 Evaluation Setup

In order to evaluate the performance of the system, and to assess its scalability, we have performed a series of tests. The metrics adopted for these tests are: a) the end-to-end (E2E) response time: the time between the issue of the request by the user and the end of reception of the response by her browser and b) the throughput of the web server (expressed in requests per minute, rpm).

The variable parameters are: a) the period of sensor checks. We test the system for 500ms, 1s, 2s and 5s periods of sensor checks and b) the number of concurrent users. We test the system for 1, 2, 5, 10, 20, and 50 concurrent users.

For each test, a number of threads equal to the number of concurrent users are created, and, subsequently, each thread requests a page from the system.

The evaluation was performed on a desktop PC with an AMD Athlon XP 2600+ CPU, 512MB of RAM, running Windows XP Professional Edition, and the Tomcat servlet container v.5.5 [14]. The client browser is the JMeter tool v2.0.3 [12], which adds virtually no overhead when compared to a typical HTML browser (e.g., Mozilla).

The client runs on a laptop with an AMD Mobile Athlon XP 1.6 CPU and 256 MB of RAM, running Windows XP Home Edition.

The source XForms document has a size of 1687 bytes. The resulting HTML file has a size of about 2100 bytes (depending on the CSS tags used each time).

The test is further divided into test groups. Each group is characterized by a different *sensor check period*. For each group, the test measures the system response time and the web server throughput for different numbers of users and for a network connection speed of 11 Mbps (WLAN). This is can be considered a *stress test*, which simulates page updates for each sensor check. In other words, this is a worst-case scenario where all sensor checks lead to page updates.

In order to reduce the statistical error, a fairly large (around 500) number of samples (i.e. total number of requests) are collected in each measurement.

Table 4. Test setup matrix

Sensor check period	0.5, 1, 2, 5 seconds
Bandwidth	11 Mbps
Requested page size	~2 Kbytes
Page refresh	At every sensor check
# of users	1-50
# of samples	~500

4.2 Evaluation Results

From the performed tests, it can be readily deduced that the longer the sensor check period is, the more scalable the system is. This is due to the (expected) fact that if the sensors are checked less frequently, fewer requests per minute are issued to the server, and, thus, the latter is capable of handling more users concurrently.

The *maximum throughput* (see Fig. 3) the system has achieved is slightly over 400 requests per minute (rpm). With a 5-second sensor check period, the system serves a single user at a rate of 12 rpm and scales well up to 40 users with 400 rpm. With lower periods, the system is saturated earlier, (20 users for 2s, 10-20 users for 1s, 5-10 users for 500ms); this is reflected, among other things, in the increased response times.

The *E2E response times* follow the same trend for the whole range of sensor check periods (see Fig. 4). While increasing with the number of users in a roughly linear fashion, that increase becomes sharp after the 20-user mark (less sharp for a 500ms sensor update period). It can be easily seen that, in general, a longer sensor check period reduces the response time.

The main conclusion is that, since the system needs roughly 130ms to process and reply to each individual request, the rate at which the requests arrive to the system should be less than that, i.e. roughly 460 rpm. This is close to the observed maximum throughput of 410 rpm. Thus, the more users the system has to support, the longer the sensor check periods must be so as to avoid server saturation and delays. For 1-5 users, the sensor check period can be as low as 500 ms, but for 50 concurrent users, the period must be at least 5s.

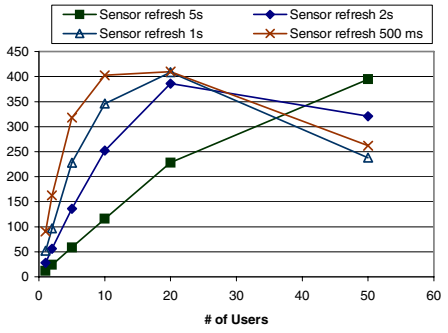


Fig. 3. Server Throughput (requests/min)

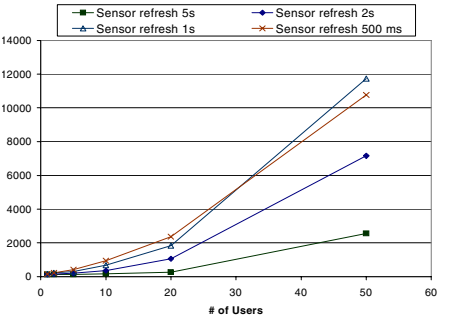


Fig. 4. Average E2E Response Time (msec)

5 Related Work

The presented system, although supporting multimodality of user interfaces, focuses on *presentation adaptation*. To our knowledge, this is a new concept in adaptive systems. In fact, we are not aware of any work sharing our objectives. The most relevant research fields are content adaptation, based either on device capabilities or user context, and multimodal interfaces. In the following paragraphs we briefly present such efforts.

The field of adaptive and context-aware systems is currently undergoing intensive research. However, most of these research efforts are focusing on the adaptation of the served content itself, rather than of its presentation. Examples of content adaptation can be found in [1], [2], and [5].

Additionally, there is active research in multimodal systems, which transform an interface written in a generalized, abstract language to a concrete user interface language, appropriate for the client device. Indicative tools in this field are [3] and [13]. Another proposal, presented in [6], describes a system that selects the categories of the content that will be presented to the client (text, pictures or sound), based on the capabilities of her device.

6 Conclusions

In this paper we have presented Chameleon, an innovative system for dynamically adapting Web user interfaces upon changes in the user's environmental conditions. Key enabler for the system is the wireless sensor network that allows the system functionality to be provided to mobile users. The implementation of the system with established technologies and the performed evaluation showed that the concept of "context-aware presentation" is quite promising and could be incorporated in next generation products and services. In fact, as the sensing technologies advance and are embedded in end-user devices, we expect more research efforts towards this direction. Such efforts should also address user evaluation issues as well as develop a theoretical basis for context-driven presentation adaptation. For instance, the required expressiveness of the adaptation rules is still an open issue. Research in uncertainty

reasoning, fuzzy logic and defeasible reasoning may provide some useful modeling and reasoning tools for designing more advanced system behaviors.

Among our planned research tasks, are the integration of multimodal support (i.e., target languages such as WML and VoiceXML), apart from the currently supported HTML, and the evaluation of the system with real users.

Acknowledgement. This work was partially supported by the Greek General Secretariat for Research and Technology (GSRT) under the framework PENED 2003 (research grant 03ED173).

References

[1] Billsus, D., Brunk, C.A., Evans, C., Gladish, B., Pazzani, M.: Adaptive Interfaces for Ubiquitous Web Access. Communications of the ACM 45(5), 34–38 (2002)
[2] Cheverst, K., Mitchell, K., Davies, N.: The role of Adaptive Hypermedia in a Context-Aware Tourist GUIDE. Communications of the ACM 45(5), 47–51 (2002)
[3] Mayora-Ibarra, O., Cambranes-Martínez, E., de la Paz-Arroyo, O., Fuentes-Penna, A.: A Visual Programming Environment for Device Independent Generation of User Interfaces. In: Proc. 1st Latin American Conf. on Human-Computer interaction, Brazil (2003)
[4] Abrams, M., Phanouriou, C., Batongbacal, A., Williams, S., Shuster, J.: UIML: An Appliance-Independent XML User Interface Language. WWW'8, Toronto (1999)
[5] Samulowitz, M., Michahelles, F., Linnhoff-Popien, C.: Adaptive Interaction for Enabling Pervasive Services. In: Proc. 2nd ACM Int'l workshop Data Eng. for wireless and mobile access (MobiDE), Santa Barbara, CA., U.S.A (2001)
[6] Yew, A., Liotta, A., Yang, K.: Adaptive Content for the Mobile User: A Policy-Based Approach. In: 2nd Int'l Workshop on Mobility Aware Technologies and Applications, Montreal, Canada (2005)
[7] Schiller, J., Voisard, A.: Location-Based Services. Morgan Kauffman, San Francisco (2004)
[8] Anagnostopoulos, C., Tsounis, A., Hadjiefthymiades, S.: Context Awareness in Mobile Computing: A Survey. Mobile and Ubiquitous Information Access Workshop, UK (2004)
[9] Crossbow. Wireless Sensor Networks http://www.xbow.com
[10] Wugofski, T. et al.: CSS Mobile Profile 1.0. W3C Candidate Recommendation (2002)
[11] Meyer, E.A., Bos, B.: Introduction to CSS3, W3C Working Draft (2001)
[12] Apache Software Foundation: Apache JMeter http://jakarta.apache.org/jmeter/
[13] Harmonia Inc: LiquidUI Tool http://www.harmonia.com/index.html
[14] Apache Software Foundation: Apache Tomcat http://tomcat.apache.org/
[15] Dubinko, M., Klotz, L., Merrick, T., Raman, T.V.: XForms 1.0, W3C Recommendation (2003)
[16] Clark, J.: XSL Transformations (XSLT), Version 1.0. W3C Recommendation (1999)
[17] NCITS V2 - Standards Development Committee on Information Technology Access Interfaces www.ncits.org/tc_home/v2htm/V2docs/v201011.htm
[18] Zakas, N.C., McPeak, J., Fawcett, J.: Professional Ajax. Wrox Press (2006)
[19] Puerta, A., Eisenstein, J.: XIML: a common representation for interaction data. In: 7th international conference on Intelligent User Interfaces, ACM Press, New York (2002)
[20] Hill, J., Szewczyk, R., Woo, A., Hollar, S., Culler, D., Pister, K.: System architecture directions for network sensors. ASPLOS (2000)

Effectiveness of Communication Process Support for Senior Citizens with Information Machines

Yosoko Nishizawa[1], Yoshinori Wakizaka[1], Takayuki Yamamoto[2], and Harumi Hasebe [3]

[1] TOSHIBA Corporation, Design Center,
1-1, shibaura 1-chome, Minatoku, Tokyo 105-8001, Japan
[2] Toshiba HA Products Co., Ltd., Development And Planning Group,
1-6 Ota Toshiba-cho, Ibaraki, Osaka 567-0013, Japan
[3] Toshiba Consumer Marketing Corp, Business Strategy & Planning Group,
2-2-15 Sotokanda, Chiyoda-ku, Tokyo 101-0021, Japan
{Yosoko Nishizawa,Yoshinori Wakizaka,Takayuki Yamamoto,
Harumi Hasebe}hcdhci@design.toshiba.co.jp

Abstract. Senior citizens using IT services have two types of barriers; difficulty of machine use and lack of motivation to use service. External support is necessary to lower these two barriers by enhancing interface usability and promotion of use. In this paper, experiments that were performed to prove the effectiveness of such supports were discussed. The results of the experiment showed that promotion of communication among users was effective to lower both of the two barriers.

Keywords: senior citizen, communication, information machines, IPv6.

1 Introduction

Some machines for IT such as PCs are not yet in common use by senior citizens aged 65 and older. In Japan, the rate of PC usage is 22.7% among those in their late 60s and 7.9% for those in their 70s. It is indicating that a large generation gap remains[1]. As the cause, the existence of the necessity of using, and the difficulty of an operation of the equipment itself can be considered. In a similar previous study, these were the top two reasons given for not using the Internet[2]. Past research reports also suggested that there were psychological barriers to the use of information machines by senior citizens and that, in addition to technical support, senior citizens require content that would increase their eagerness to use the information machines[3],[4].

Considering this state of affairs, we can easily imagine that even though there are services useful to senior citizens, there are many situations in which senior citizens cannot take advantage of them simply because they have to go through information machines. In the future, providers of services need to consider not only how to provide better services but also how to make services accessible to more users.

For this report, Toshiba worked with the city of Yamanashi and the University of Yamanashi to use next-generation Internet technology (IPv6: Internet Protocol version 6) in an experiment running from December 2003 to February 2004[5]. Based

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 416–425, 2007.
© Springer-Verlag Berlin Heidelberg 2007

on the experimental data, the research clarified the usage status of information machines among senior citizens and the barriers to such usage. The research also confirmed the effectiveness of usage promoting mechanisms put in place to remove those barriers.

2 Senior Citizens' Inclination to Use IT Services

We interviewed senior citizens to find out about their lifestyles and how they were using information machines. Based on the results, we discuss senior citizens' perceptions of information machines and the possibilities for incorporating information machines into their communication activities.

2.1 IT Services in the View of Senior Citizens

Interviews were held with three healthy senior citizens, male and female, in their 80s and residents of the city of Yamanashi. Constructive interviews were held at the subjects' homes. The interviews focused on three general topics: 1) the subject's current situation (public services used, communication, degree of contact with information); 2) needs (challenges with their current situation, what they would like to do); and 3) hypothetical surveying (possibility of accepting information services, likelihood of using information machines).

The result indicates that the senior citizens were more eager than expected to actively acquire information. It also revealed, however, that although the subjects wanted to use information machines, they were not doing so or could not do so presently because it seemed difficult and they did not know how. The most important fact revealed was that the subjects were satisfied with their current machines and their activities and experience. So although they were interested in new machine, they did not feel much need for using them.

2.2 Barriers to IT Use by Senior Citizens

For senior citizens, using IT services is a new experience in which they use machines they have never used before. There are two problems before senior citizens use IT services: lack of motivation to use service and difficulty of machine use.

The barrier concerning lack of motivation to use service is that although they understand logically that using these services would add to their convenience, they are actually hesitant to try, because senior citizens are satisfied with the status quo. They already use without problems, such as ATMs and cell phones, etc. Once the user tries using these things, they continue to use them, and do so without hesitation like these apparatus. For this reason, a sense of unease toward new experiences as well as the sense that they would be too much trouble appears to be the main barriers. If a person tries something once and can be persuaded that this new experience is convenient for him or her, it opens the possibility of continuing use.

The barrier concerning difficulty of machine use is expressed in two ways. In the first case, the user does not know how to use a machine and does not touch it for fear it will break. In the second case, the user uses the machine, but finds that operation is

complicated, so they prematurely give up using it. The fear of breaking a machine and the irritation by difficulty in using process, are thought to be the reasons the user give up. For these reasons, one issue is how to design machines so they are easier to use.

The two barriers are really opposite sides of the same coin. During initial use of a machine, the user has to work hard learning how to use it. On the other hand, as the senior citizen learns to operate the machine and becomes able to accomplish his or her intended function, the experience derived becomes accepted as something beneficial. This beneficial experience leads in turn to further use of the machine. If the process of learning to operate the machine does not go well, however, and it is too difficult to achieve the intended function, then the person will feel that the experience of using the machine or service was uncomfortable. In that event, continued use becomes mentally taxing and the person does not gain experience.

It is predicted that systems and machine designs allowing senior citizens to appreciate the benefits of experience and learning would mitigate the barriers and expand the usage possibilities.

3 Overview of Experiment

In this section we verify whether mechanisms and machine designs implemented lowering of the barriers actually encourage senior citizens to use IT services.

3.1 Experimental Machines and Participants

The network used in the experiment was that of Yamanashi CATV Ltd. It linked the subjects' homes with the Yamanashi City Hall and the testing office. Figure 1 shows an example of the terminal in use.

Fig. 1. The IT home terminal. The terminals used by the subjects were equipped with teleconferencing camera and microphone and were controlled by a touch-sensitive panel.

The subjects were recruited from 30 households housing a person of 59-82 years of age living in the city of Yamanashi. Of these, about two-thirds of the participants claimed they had no experience with information machines.

3.2 Systems and Services Provided for Verification of Experiment

Six services whose senior citizen probably gets interested were provided. The notification service, teleconferencing function, calendar and bulletin board could only be used among the subjects. However the subjects could use the e-mail and Internet services also with non-subjects.

The machine could be operated by a touch panel, which was usable even by those unaccustomed to keyboards. The user interface featured large characters and buttons, and screens were developed based on the results of the pre-survey and designed to be operated as easily as possible by senior citizens.

3.3 Mechanisms for Promoting Usage

We intervened to control the experiment rather than giving assignments full discretion. This intervention consisted of various conditions and mechanisms with the focus on technical support and presentation of assignments during the period of the experiment. Figure 2 gives the experiment schedule and a list of steps taken.

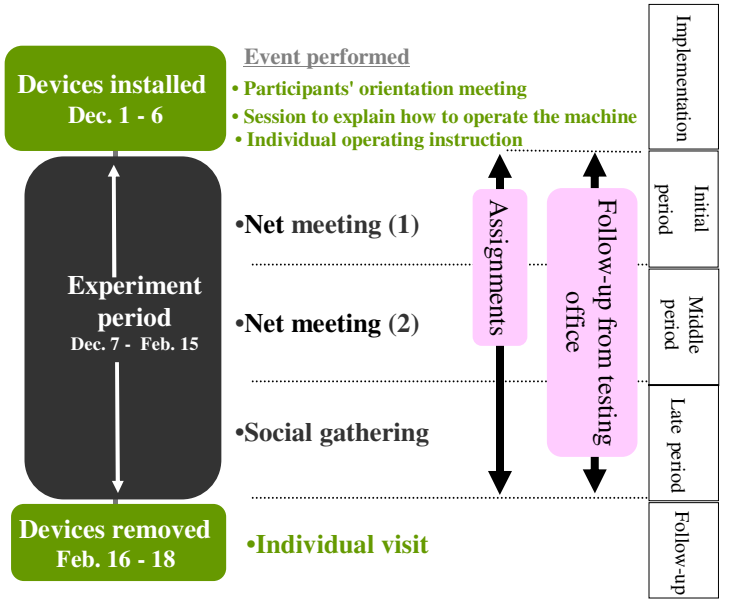


Fig. 2. Experimental mechanisms. Various technical support and assignments were given during the experiment to the experimenter.

The follow-up from the testing office consisted of appeals, by e-mail and notification, based on promoting the use of machines and services intended to promote communication with participants outside the group that applied. An advice center was also set up to deal with any problems encountered while using the machines, so there was constant operating support.

Assignments were given that used teleconferencing, e-mail and the bulletin board in order to encourage communication within and beyond the group applying to be in the experiment.

Miscellaneous events were held to create motivation to participate in the experiment and lower the barriers to operating the machines. Meetings were held to orient the subjects and explain the machines at the time of implementation, and subjects also received individual explanation in the home. At events held during the experiment period, subjects had a net meeting with the city mayor by teleconference, and subjects teleconferenced to discuss their opinions. A teleconferencing week was also set up, during which subjects held teleconferences at set times.

Assignments were given to be performed, however, no penalties or other psychological pressure was used. We asked subjects to complete these assignments, and the subject can use the system voluntarily.

3.4 Actual Use (Experiment Results)

Figures 3 show the results for the two months of the experiment.

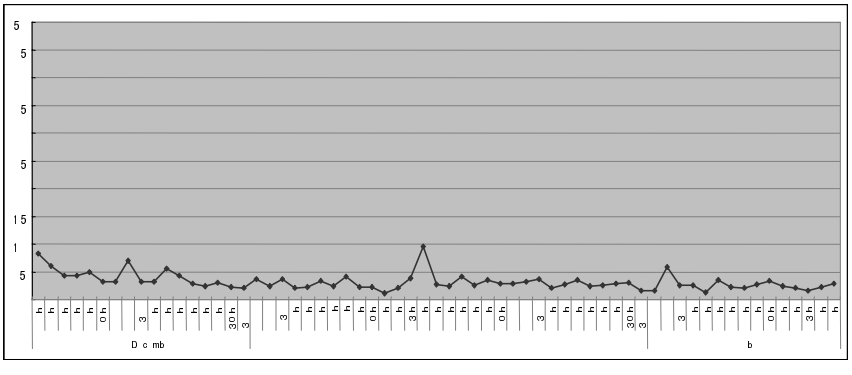


Fig. 3. Machine usage frequency. The vertical axis of the graph shows the number of times the machine was turned on, while the horizontal axis represents the time of the experiment (from December 15, 2003 to February 15, 2004).

Figure 3 shows the average number of times the machine was turned on for all subjects. The average daily frequency for all participants was three to four times a day throughout the experiment period. The data shows that there was not any one day in which none of the subjects turned on the experimental machine. The high number of uses at the beginning of the experiment (in mid-December) was influenced by the orientation sessions held at the users' homes when the experimental machines were being installed. Users on average did not subsequently greatly reduce their frequency of use. That suggests the experimental machines found steady use.

The month-by-month usage frequencies for the services show a declining trend from the start to the second half of the experiment, but nearly all the services were used once a day. The bulletin board and teleconferencing function were particularly

well used. The results of the post-experiment questionnaire show that at least 50% of the user used all the provided services, with the percentages of user using the bulletin board and teleconferencing function being particularly high. The subject recognized it as having used well.

The above results show that the senior citizens were able to use machines and services previously unfamiliar to them.

4 Acceptance of IT Services by Senior Citizens

Below is a discussion on the effectiveness of external support for encouraging acceptance of IT services.

4.1 Effectiveness of Technical Support

To avoid a situation where subjects initially refused to use the services because they did not know how to operate them, we provided three types of training as technical support: a session to explain the experiment; a session to explain how to operate the machine; and individual operating instruction.

An advice center was also available during the period of the experiment. The center provided support by telephone, and in those cases where the telephone was insufficient to deal with the situation, the testing office provided operating support when its personnel met with the subject, such as when sending someone out to teach the subject or when taking surveys.

Most of the subjects in this experiment were beginners at using information and communication machines. For that reason, it was predicted that in the initial stages the subjects would feel unsure whether they would be able to operate the machines, that they might not adequately understand the machines, and that these would be psychological barriers to accepting them. From the experiment result, it appears that the devices were well used overall. On a questionnaire given at the session for explaining the machine (before the experiment), more than 90% of the respondents said they thought using the machines would be fun, however about 80% said they felt confused about the machines. The advice center, moreover, was used nearly every day throughout the experiment period. During the first half of the experiment, there were at least 10 inquiries per day, which declined to just a few per day in the second half. In the final 10 days before the experiment ended, it would not have been a problem to shut down the advice center.

This indicates that the various training sessions held before the experiment, although they may have implanted an inclination to use the machines, did not adequately eliminate the subjects' sense of unease about operating machines. Also, nearly all the subjects felt unsure while operating the machines and inquired at the advice center. In interviews taken after the experiment, many of the subjects said that getting advice by telephone, coming in addition to individual machine explanation, was very effective.

These forms of support therefore apparently helped the subjects feel at ease using the machines.

4.2 Information and Communication Machine Acceptance and Usability

Results of the post-experiment questionnaire on user friendliness of the interface are given in Figs. 4. Positive responses on the question of ease of use were lower even than those concerning the impression of the experimental machine. The reason is that there are significant problems of usability, such as character input.

The assessment of user friendliness of individual functions was low for character input. This is related to the operation of some major functions, the e-mail and bulletin board, so these appear to be important interfaces for users. A low assessment was also given to the microphone and speakers, items related to senior citizens' physical abilities.

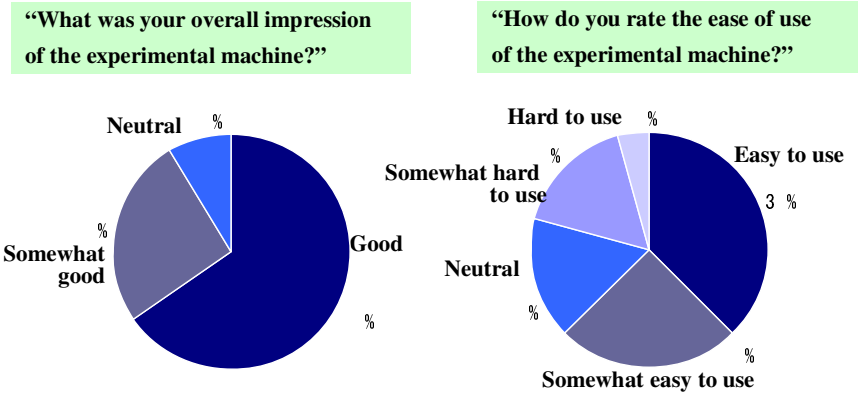


Fig. 4. Impression of machine and user-friendliness. The left graph a left graph is the answer to the question of "What was your overall impression", the right graph is an answer to the question of "How do you rate the ease of use".

Additionally, although a hiragana-only keyboard was provided from the start, users asked for a character conversion function as they used the machines more, so a standard soft keyboard such as is commonly used was provided. Furthermore, the inquiries to the advice center shifted from the very basic (such as how to turn the power on and off) to the complex (how to do things that were not necessary in standard use). This suggests that as the subjects used the machines more, the scope of that use increased and they started to use advanced functions. The post-experiment interviews yielded many comments supporting these views.

The above suggests that the level of the subjects' operating proficiency rose over time, as illustrated in Fig. 5.

Many of the subjects in this experiment were beginners at using information and communication machines. In the initial stage of introducing the machine, the subjects felt unsure of whether they would be able to operate the machines as well as not adequately understanding the machines. These beliefs became psychological barriers to accepting the machines. Performing assignments, moreover, helped them grow accustomed to operating the machine to some extent, and to overcome the barriers to learning how to operate it. Subsequently, as time passed, the user became more

proficient, and when they had improved a certain amount, they had reached the limit of skill improvement.

It was also shown that important factors for lowering the learning barriers were an easy-to-use interface and services that were themselves easy to use, as well as reducing machine problems. In any case, a necessary condition for senior citizens to accept a machine is the ability to use the machine and accomplish the intended task, even if the user has just begun using the machine. Machines must be designed to meet this condition.

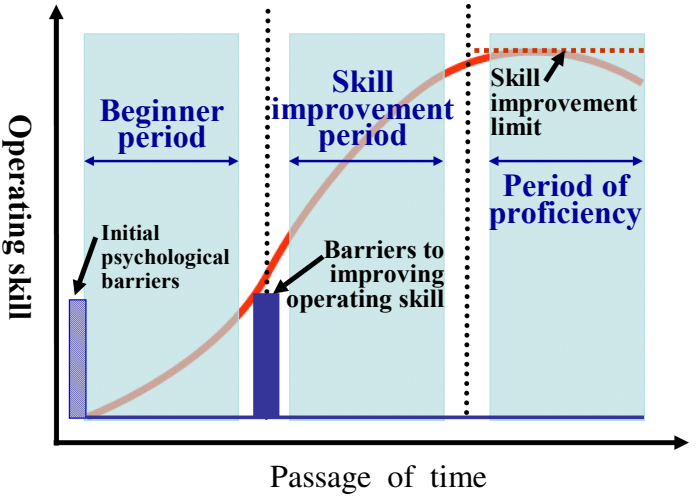


Fig. 5. Proficiency over time. This figure is expressing the mastery process in apparatus exploitation typically.

4.3 Effectiveness of Supports for Promoting Usage

During the experimental period, using of the system was performed by a subject's free will. In order to promote communication, however, subjects were given such assignments as sending New Year's greetings to each other by e-mail and posting notes to the bulletin board. Opportunities to use the machine were also given, such as a net meeting with the city mayor using the teleconferencing function. To help the subjects browse information frequently, the testing office frequently updated the information displayed on the screen and sent messages and greetings by e-mail every day.

Only about one-third of the subjects actively carried out the assignments of sending e-mail or posting messages on the bulletin board, whereas more than 80% of the user read about them. This indicates that giving assignments itself was effective for some subjects, but did not lead directly to usage of the machine. The fact that subjects were browsing the information, however, suggests that their perception of participating in the experiment and the sending of e-mail from the testing office made the subjects feel obligated to check the content of the messages, which promoted usage.

Figures 6 show the frequency of usage of each service for two subjects with very different levels of usage frequency.

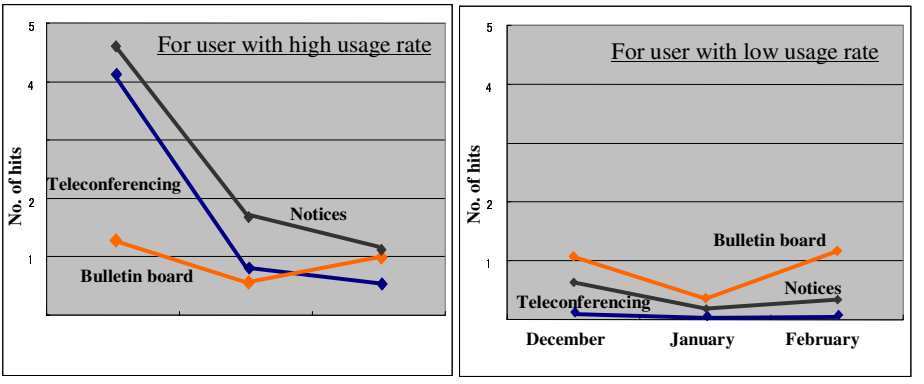


Fig. 6. Service usage frequency .The left graph expresses the number of times of exploitation of the moon average for every service of a user with high exploitation frequency, the right graph expresses for every service of a user with low exploitation frequency.

The trends in the frequency of using the teleconferencing and notification functions show that the subject with a low usage rate started out with low usage and this did not change greatly, but the subject with high usage rate greatly reduced the number of uses in the second half as compared to the early period. This subject had been looking forward to using the teleconferencing function since before the start of the experiment, and in the early stages had tried using it frequently. However, after this person tried dialing a teleconference several times and the other party did not answer, the subject gave up after a few weeks. This situation is also seen in e-mail and other services. In post-experiment interviews, subjects said that they quit using a service because they were not getting replies.

In contrast, the bulletin board suddenly got increased use in February, for both the subject with high usage and the one with low usage. The subject with low usage in particular used this service more often than in the early stages. This is because a social gathering was held to build friendships among users in the latter half of January. Up to that point, the users' only exchanges had been over the network. So meeting actually was very effective in promoting usage. The subjects made active use of the machine at the social gathering without any encouragement by the testing office. There was active use of the bulletin board. By using this service, subjects did not require a specific partner for interaction but could have communication with unspecified others. Looking at the number of messages posted by each subject, there were very active users and less active users, however even less active users used the service, participating as readers only. There were many users who enjoyed reading messages on the bulletin board.

The above indicates that a sense of obligation and achievement of an objective are important factors for promoting usage, and it has also been shown that community-building is also significantly related.

5 Support for Promoting Use by Senior Citizens, and Issues

The experiment found two types of psychological barrier to the use of information machines by senior citizens: difficulty of machine use and lack of motivation to use service. The experiment further showed that the characteristics of each of these vary from the initial usage stage (acceptance) and the stage of the usage process.

The difficulty of machine use is uneasiness and lack of understanding about the machines. As the person continues to use the machines, it is necessary to expand the functionality to meet the user's requirements. And it is basic to this need that the purpose is attained, when using. Meeting this need requires sufficient follow-up to eliminate the sense of unease, and it is also necessary that the machines can respond flexibly in keeping up with the user's level of proficiency.

The lack of motivation to use service includes a sense of unease toward new experiences and a sense that they are too much trouble. For overcoming this, it is required applying inducement to try the experience and helping the subjects understand the merits of the experience. Communication is effective for helping user to understand the merits, and having systems in place to encourage community-building lowers the barriers relating to inclination to use services.

Technical problems are closely related to problems with users' motivation to use the machine, so both of these have to be solved simultaneously to encourage usage. Many issues, such as the varying effectiveness of different approaches to support, remain. We believe that illuminating these issues will allow more people to take advantage of the benefits of IT services.

References

1. Ministry of Internal Affairs and Communications, Communications Usage Trends Survey Results (Attachment) (2005)
2. Ministry of Internal Affairs and Communications, FY2002, Communications Usage Trends Survey Report (2002)
3. Iida: Aging Societies and Regional Information Systems. In: Abstracts of the Second Conference of the Japan Society for Socio-Information Studies, pp. 48–49 (1997)
4. Komatsu, Ishihara, et al: Changing the Psychological Barriers and Motivation for Senior Citizens to Use Networks, Papers from the Cell Phone and Car Navigation Usability and Ergonomics Symposium, pp. 103–108 (2002)
5. Toshiba Corporation, Report on R&D and Proving Tests of Application Software for Silver Net Society with IPv6 Information Appliances

Technology Advances and Standardization Toward Accessible Business Graphics

Hironobu Takagi and Tatsuya Ishihara

Tokyo Research Laboratory, IBM Japan, 1623-14, Shimo-tsuruma, Yamato, Kanagawa, Japan, 242-8502
{takagih,tisihara}@jp.ibm.com

Abstract. Various types of graphics are exchanged in our daily business and education processes. In spite of the importance of business graphics, they are not accessible for visually impaired people, especially for the blind, and this impacts their productivity at work. Current serialization-based screen reading techniques do not provide sufficient functionality for accessing graphics. In this paper, we report the results of our survey to identify next generation accessibility features in future graphics standards, especially for the OpenDocument Format (ODF). We will first compare accessibility functions in various types of existing standards. Then we will report our survey results for three related areas, a survey of existing business graphics in presentation documents to unveil the complexity of practical graphics, a survey of research on various types of graphical taxonomies, and a survey of interface technologies for representing graphics non-visually, such as screen reading and pictorial Braille. Finally, we will propose three practical enhancements for standard graphic formats based on the survey results.

Keywords: Accessibility, graphics, presentation, visually impaired, standards, OpenDocument Format.

1 Introduction

Various types of graphics are exchanged in our daily business and education processes, such as organizational charts, process diagrams, timeline charts, software architecture charts, mathematical and statistical charts, and so on. These business graphics convey vital information and knowledge about business projects, and they should be shared among people who work together on or in the projects.

In spite of the importance of business graphics, they are not accessible for visually impaired people, and thus impact on their work productivity. This barrier is becoming severe as business visualization extends deeper into business processes. Currently, no standardized conversion method from visual graphics into non-visual media is known, since the bandwidth of vision is incomparably wider than any non-visual senses, such as the senses of hearing or touch.

Currently, screen readers and voice browsers are the major methods to access arbitrary documents, but they do not have sufficient functionality for accessing graphics. Screen reading is a method to read aloud (or display one line of Braille on a refreshable Braille device) focused content fragments according to users' keyboard

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 426–435, 2007.
© Springer-Verlag Berlin Heidelberg 2007

operations. Blind users are required to recognize the underlying visual structure by exploring among the graphical objects by using a keyboard. In other words, the graphics are serialized into a sequence of text fragments, and users need to imagine the corresponding visual structures through a keyboard-based exploration process instead of eye-movement-based visual exploration.

As members of the OASIS OpenDocument Format (ODF) accessibility sub-committee, we have been working to improve the accessibility of standardized office document specifications including those for presentation documents. The presentation document specification in the latest ODF (version 1.1) has basic accessibility functions, such as alternative text for a group made of graphical primitives, graphics-caption relationship assignment, and reordering functions for tab navigation. These functions are comparable to any other vector graphics format, but it is clear that these functions are insufficient to make most business graphics accessible.

In this paper, we report on our survey to help design next generation accessibility features in the future versions of the ODF presentation documents specifications. First we will compare accessibility functions in various types of existing standards. Then we will report our survey results for three related areas, research on graphical taxonomies, existing business graphics, and interface technologies. Finally, practical future improvements to the standards will be discussed.

2 Existing Accessibility Standards for Business Graphics

Table 1 shows a comparison of existing standards for documents and graphics, and their considerations regarding accessibility, especially for the blind. We investigated five content standards, W3C HTML 4.1, W3C/OASIS WebCGM 2.0, W3C SVG (Scalable Vector Graphics) 1.2, ECMA Office Open XML 1.0, and OASIS OpenDocument Format (ODF) 1.1. A draft metadata standard, W3C ARIA, is also shown in the table for comparison. It is designed to complement W3C content types (HTML, WebCGM, and SVG) in terms of dynamic Web content accessibility.

We classified accessibility functions into seven categories, five categories for text descriptions, and one category each for object groupings and navigation orders. In other words, the union of the current accessibility functions in the various standards only covers these concepts for accessible graphics. The goal of this paper is to identify any missing concepts and to investigate future possibilities for enhancements.

“Text descriptions” are the basic method to make graphics accessible for screen readers by adding text alternatives. Most of content types have functions to add at least two levels of invisible text descriptions (short and long) to each graphical object. “Speaker notes” are not officially defined for accessibility purposes, but are widely used to describe presentation slides. Only ODF has a function to associate a visible text with a graphical object as a description (“Visible label”). HTML, SVG, and WebCGM will have association functions, since the ARIA draft has two types of metadata for the association (describedby and labelledby).

“Object groupings” allow content authors to mark a set of graphic elements as a semantically indivisible object and to add text descriptions to the group of objects. This function can also be used to create hierarchical structures by adding nested groupings. All vector graphics types have grouping functions, and ARIA defines several types of object groupings.

Since screen-reading is the basic method for accessing graphics, each content type has functions to compensate for the order of screen-reading ("Navigation orders"). In order to mark the orderings, we found three types of methods for ordering objects. ARIA's "tabindex" is an attribute to assign a priority value to each object. It is easy to adjust the reading orders among a small number of objects, but it tends to be difficult to control large numbers of objects. ARIA's "flowto" is a method to point to a next object to set a relevance-based reading order, but improper assignments can cause inappropriate skipping of content. ODF's "draw:nav-order" is a method to describe a sequence of object IDs in a property list. This method is simple and easy to manage, but no interface to change this attribute has yet been implemented in any office document editor. Therefore, though the idea seems promising, it is still too early to assess its utility.

Table 1. Comparison of existing graphic standards in terms of accessibility for the blind

			W3C HTML4.1	W3C/OASIS WebCGM 2.0	W3C SVG 1.1	ECMA Office Open XML 1.0	OASIS ODF1.1	W3C-ARIA (draft)
Graphics type			images	images and vector graphics	images and vector graphics	images and vector graphics	images and vector graphics	
Attributes / Elements for Accessibility	Text descriptions	Short descriptions	ALT	content	title	alt(VML), descr(DrawingML)	svg:title	
		Long descriptions	LONGDESC	layerdesc	desc		svg:desc	
		Regional descriptions	MAP & ALT				draw:image-map & svg:desc	
		Speaker notes	NA	NA	NA	p:notes	presentation:notes	
		Vsible labels					draw:caption-id	described by, labelledby
	Objects groupings		NA	grnode	g	p:grpSp	draw:g	group, region, section, list, etc.
	Navigation orders		NA				draw:nav-order	tabindex, posinset, flowto, etc.

3 Survey

This survey aims at identifying important missing functions in the current standards. Three types of different surveys were conducted to cover a wide range of related knowledge, and to anticipate future accessibility environments.

3.1 Survey of Existing Business Graphics

The first approach was an analysis of existing business graphics. We collected more than 4,000 pages of presentations from the Internet and also various types of graphics used in our company, and then selected about 60 pages as typical examples. Fig. 2 shows an overview of these charts. The horizontal axis shows the complexity of the graphics and the vertical axis shows the formality. We loosely categorized these graphics into four categories. “Simple graphics” are relatively simple, so we think that these graphics can be translated into accessible formats in various ways, such as pictorial Braille or alternative voice access interfaces (see Section 3.3).

“Formal diagrams” are formally defined diagrams, which have a well-defined data-model behind the visual structure, such as tree structures, Venn diagrams, cause-and-effect diagrams, and UML (Unified Modeling Language) diagrams. If a non-visual conversion were defined for each data model and blind users learned the appropriate access method, then these diagrams should be accessible. In order to implement this approach, the graphics should have functions to preserve the data-model for each part of the figure.

“Table structures” are graphics that can be regarded as tables. If the visual information of a graphic can be represented as a table, then blind users can access them by using table navigation functions in screen readers. In this case, standards would need to have functions to add tabular metadata to the graphical objects.

“Composite diagrams and graphics” have multiple types of diagrams in one figure, annotated formal graphics, or visually represent multiple aspects (dimensions) of

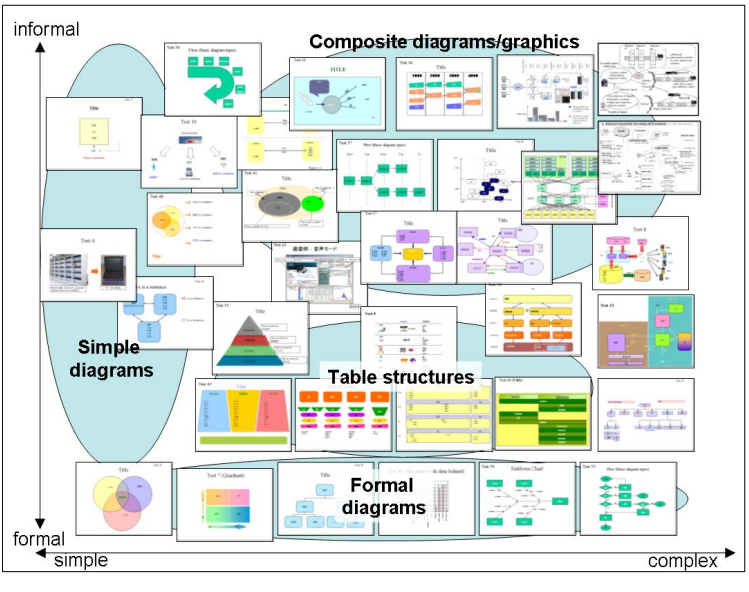


Fig. 1. Variety of collected existing graphics in presentation documents

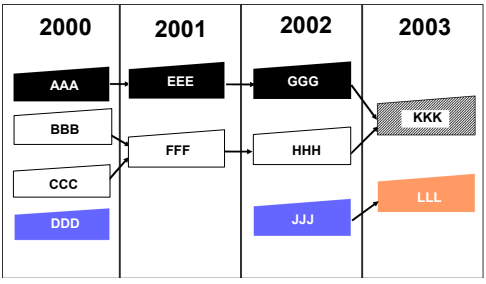


Fig. 2. Example of a graphic with multiple aspects. The color scheme indicates categories of topics (nodes), the horizontal positions show a timeline, and the arrows indicate relationships between the topics.

some information in one figure. In the case of Test 46 (Fig. 2), the basic diagram structure shows a history of topic changes and each vertical box indicates a year.These two aspects are well represented in the diagram, but it is difficult to represent these aspects non-visually. At this time, general methods to make them accessible are not well known. In fact, more than 50% of the investigated graphics were classified into this category.

3.2 Survey for Graphical Taxonomies

That survey of existing graphics showed us the complex nature of practical graphics. It raised questions about the possibilities of making them accessible, since they contained various graphical objects with various visual properties and they were aligned to express specific knowledge and information. In this survey, we will try to figure out possibilities to describe complex graphics by referring to research on graphical taxonomies.

Lohse et al. [1] classified various types of graphics into eleven categories based on the results of user studies, such as “structure diagrams”, “cartograms”, “maps”, “graphic tables”, “process diagrams”, “icons”, “time charts”, “network charts”, “pictures”, “tables”, and “graphs”. Their coverage of graphics types is larger than our survey as described in the previous section, since they collected a variety of graphics from popular reference books on graphics. Their classification and user studies show us the breadth of the problem space here.

Blackwell and Engelhardt [2] conducted the most comprehensive survey for diagram research, and it resulted in a set of meta-taxonomies. This meta-taxonomy yields a useful framework for the necessary metadata in presentation documents. Among their nine taxonomic aspects, at least six representation-related aspects have a close relationship with metadata for accessibility, from “graphic primitive elements and the graphic properties” to “represented information”. Between “graphic primitive” and “represented information”, two aspects for symbols (“conventional elements” and “pictorial abstraction”) are defined, and “graphic structure” is an aspect for the two-dimensional structure of graphics, such as linear sequence, two-axis-chart, table, or tree structure. The “mode of correspondence” is a unique aspect, which

describes the relationship between a representation and its meaning, such as literal vs. metaphorical, direct vs. indirect, and iconic vs. symbolic.

It is clear that the "represented information" is the most important information to be presented non-visually, however no standards support bundling of the represented information in graphics. The idea of "data-model attachment" is an important idea (See Section 4). It is not clear whether "mode of correspondence" can help non-visual access or not, so we need to investigate this aspect further. The "graphic structure" should also be one of the focuses of future standardization, since it helps translate into non-visual formats. The aspects for symbols are handled well in current standards as various levels of text description. The "graphic primitives" are less important, since they are already preserved in the file formats as objects and properties.

Kong et al. [3] conducted a survey of low-level taxonomies for graphic structures. Since they aimed at defining a formal graphical grammar for their visual language system, the taxonomy only includes metadata suitable for automatic detection, such as topology (such as "touch" and "overlap"), direction (an 8-point compass), distance (close, far) and alignment (such as "gap", "middle", "left" and "right"). The SVG RDF proposal [4] also includes such low-level metadata, such as topology (such as "Contains", "isPartOf" and "InFrontOf"), direction (InPosition, InDirection), and alignment (AtLeft, AtRight, OnTop, Under). The necessity of this low-level metadata for accessibility purposes needs to be considered, since it would impose heavy work requirements on the creators of graphics, and the benefits for accessibility are not clear.

3.3 Survey for Interface Technologies

The third survey is of the technologies of media conversion for sensory substitution. This survey lists conversion methods that are accessed by two different human sensory systems, screen reading for the sense of hearing and pictorial Braille for the sense of touch.

3.3.1 Screen Reading

The one-dimensional speech interface is widely used and speech can even describe structural metadata information with interactions. For example, table navigation allows users to recognize two-dimensional table structures by using various commands, such as directional movement (N, E, S, W), table header reading, and consistent location reading when a user changes column or row. However, it is clear that the accessible structures are limited to relatively simple structures, and insufficient for the complexity of general business graphics.

As a trial to improve screen reader navigation, we developed a presentation access tool optimized for screen reader users, which has automatic transformation functions. We call this interface "DocExplorer" [5]. We implemented automatic detection functions for basic metadata taxonomic items, such as nodes, arrows/arcs, and groups, by analyzing primitive graphical elements. Then the tool creates a tree-view structure, which is a popular GUI control to represent hierarchical data (and familiar to many blind users).

Through the development of the automatic detection algorithm, we found various problems in improving the detection accuracy, even for such basic metadata. The

accuracy could be greatly improved (or degraded) by small graphical changes. For example, automatic detection of topological metadata for grouping, such as “in”, “crossing”, or “touching”, was affected by the precise positioning of the shapes. These results indicated a need to manually create the low-level metadata (See Section 3.2), but the workload of authoring needs to be taken into account.

We also evaluated our tool with several blind employees in our company. Considering their prior experiences, they evaluated the tool as the best text conversion tool for presentation documents. However, the reliability of the depicted structures was raised as a problem. We believe this result indicates limitations of the current technology, and also the necessity for new meta-data standards.

3.3.2 Pictorial Braille

The two dimensional tactile interface is also a common method used for reading graphical content. Today, Braille printers are commercially available, and they are in widespread use in many facilities. However, the resolution of Braille pictures is quite low compared to the original graphic pictures, because of the limitation of the human sense of touch. While the resolution of a typical laser printer is 600 dpi, the resolution of a Braille picture is about 20 dpi. For this reason, a manual conversion process is required to create Braille pictures. It takes a lot of empirical knowledge to convert graphics to Braille pictures [6].

Currently great efforts are being made to improve the tactile interfaces, and remarkable progress is being made. One newly created interface is a touch panel which implements a coordinated presentation of a Braille picture with a computer. As the user touches a region of the Braille picture, a text label attached to it is read out by a computer. This interface is already commercially available, for example in IVEO® [7] and the Talking Tactile Tablet [8]. These are useful for understanding graphical content, including graphs, maps, and so on. Because there is a lot of graphical content in teaching materials and business presentation documents, these products are expected to be popular in schools and in business environments.

Another is the refreshable dot matrix display. This interface has been studied over the years [9]. Recently the price of this device is decreasing, making it commercially available. For example, the price of the DotView produced by KGS is about $10,000 (US) for a display size of 32 x 48 dots. It is used mainly for educational purposes, because the resolution is lower than Braille pictures.

No matter which interface is used, images rendered for the blind have much lower resolution, because the resolution of tactile perception is limited (about 1 dot/mm) [10]. That is why abstraction and simplification are necessary for creating useful Braille images [6]. Many activities are going on to make automatic conversion practical [11][12]. In order to help these manual and automatic conversions, metadata supporting the simplification should be considered in the standardization processes.

4 Toward Accessible Business Graphic Standards

Through these surveys, we can see the limitations of current technologies and standards. The path to accessible business graphics environment appears long and

difficult and various activities are required. However, we found some hints about how to move toward the goal. In this section, we would like to propose three realistic improvements as the conclusion of this survey.

4.1 Mechanisms for Attaching Data Models

The survey of existing graphics showed what kinds of graphics are commonly used, and the "formal diagrams" are among them (Section 3.1). The use of these formal diagrams is increasing in enterprises, such as UML diagrams [13] and Mindmaps. These well defined diagrams have clear data-model representations in XML or other formats. Since data-models already exist (which is why the graphics exist), no additional workload is required to attach data-models to these graphics in presentation documents. These data-models are corresponds to "represented information" in Blackwell and Engelhardts's meta-taxonomy (See Section 3.2). At this time, non-visual representations of these data-models are not well established, but the transformations for simplified user interfaces and pictorial Braille transformations would benefit from standardization (See Section 3.3).

4.2 Generalization of "Connectors"

One of the most basic graphical structures in existing graphics is a directed connection between two shapes or images, which is represented using various types of arrows. Both of the presentation formats introduced in Section 2 (Office Open XML and ODF) have functions to draw directional arcs between two graphical objects. However, connectors are used only for drawing lines in the current standards. If connectors can be assigned to any type of arrow shape or image, then the connectors would become general metadata for visual connection structures. At this time, connectors are not utilized for non-visual access. It is clear that non-visual translations both for simplified user interfaces and for pictorial Braille can be improved by connector information (See Section 3.3).

This mechanism can improve the general usability for authors, since arrow graphics could follow shape relocations automatically. We believe that this is a kind of kill-two-birds-with-one-stone solution.

4.3 Preservation of Alignment Information

In Section 3.2, we introduced considerations of low-level metadata for accessibility. We need to take into account both the benefits of this metadata and the workload for authoring. The benefits are not clear, but non-visual translations, especially for pictorial Braille may benefit, since these techniques require drastic simplifications of the graphics based on the visual structures (See Section 3.3). Therefore, we would like to propose mechanisms to preserve alignments in documents recorded as metadata. Graphic authors naturally use alignment functions to layout graphic objects. Histories of these alignment operation could be used as hints for the automatic translation.

5 Conclusion

This paper proposed practical enhancements for future business graphics standards through a set of comprehensive surveys. First, we gave an overview of current graphics standards in terms of accessibility for the blind. Even for the best formats, accessibility considerations are limited by the functions of existing screen readers. Then, we first reported our survey results for existing business graphics. We collected a thousand pages of presentation images and analyzed them. We were able to organize them into four categories, but it was not clear how to make accessible the graphics in the largest category, the "composite graphics". The second survey was of research on graphical taxonomies. We found well defined frameworks for graphic accessibility. Two categories of technologies were reviewed in the third survey of interface technologies. The first approach is an expansion of current screen reading techniques, but it converts visual structures into suitable non-visual structures for screen readers. The other approach is pictorial Braille, such as paper Braille or refreshable dot matrix displays. These devices are becoming inexpensive and offer improved resolutions, but still cannot be compared to visual resolutions. That is why automatic transformation is a new and crucial technology to make them practical. Finally, three enhancements to standards were proposed. Data-model attachment will allow users to explorer the "represented information" underlying the graphics. The "generalization of connectors" will benefit both sighted graphic authors and blind end users. The "preservation of alignment information" is a low-hanging-fruit approach to add richer metadata to graphics to help with automatic transformations. We hope these proposals will be polished in the standardization process and will contribute to improve the accessibility of business graphics.

References

1. Lohse, G.L., Biolsi, K., Walker, N., Rueter, H.H.: A classification of visual representations, Communications of the ACM 37(12), 36–49 (1994)
2. Blackwell, A.F., Engelhardt, Y.: A Meta-Taxonomy for Diagram Research. In: Anderson, M., Meyer, B., Olivier, P. (eds.) Diagrammatic Representation and Reasoning, pp. 47–64. Springer, Heidelberg (2002)
3. Kong, J., Zhang, K., Zeng, X.: Spatial graph grammars for graphical user interfaces. ACM Trans. Comput.-Hum. Interact, 268–307 (2006)
4. Herman, I., Dardailler, D.: SVG Linearization and Accessibility, Computer Graphics Forum, vol. 21(4) (2002)
5. Ishihara, T., Takagi, H., Itoh, T., Asakawa, C.: Analyzing visual layout for a non-visual presentation-document interface, ACM SIGACCESS, pp. 165–172 (2006)
6. Schuffelen, M.: On editing graphics for the blind, The Hague, Holland: Netherlands Library for Audio Books and Braille (2002) http://www.stanford.edu/s̃ipma/grbl0.html
7. IVEO® http://www.viewplus.com/products/touch-audio-learning/IVEO/
8. Talking Tactile Tablet http://www.touchgraphics.com/ttt.htm
9. Kawai, Y., Tomita, F.: Interactive tactile display system: a support system for the visually disabled to recognize 3D objects, ACM SIGACCESS, pp. 45–50 (1996)

10. Way, T.P., Barner, K.E.: Automatic visual to tactile translation, part I: Human factors, access methods and image manipulation. IEEE Trans. on Rehabilitation Engineering 5, 81–94 (1997)
11. Rotard, M., Knödler, S., Ertl, T.: A tactile web browser for the visually disabled, ACM HYPERTEXT, pp. 15–22 (2005)
12. Ladner, R.E., Ivory, M.Y., Rao, R., Burgstahler, S., Comden, D., Hahn, S., Renzelmann, M., Krisnandi, S., Ramasamy, M., Slabosky, B., Martin, A., Lacenski, A., Olsen, S., Groce, D.: Automating tactile graphics translation, ACM SIGACCESS, pp. 150–157 (2005)
13. Unified Modeling Language (UML), Version 2.0
http://www.omg.org/technology/documents/formal/uml.htm

Auditory Accessibility of Metadata in Books: A Design for All Approach

Dimitrios Tsonos, Gerasimos Xydas, and Georgios Kouroupetroglou

University of Athens,
Departement of Informatics and Telecommunications,
Panepistimiopolis, Ilisia, GR-15784 Athens, Greece
{ea02534,gxydas,koupe}@di.uoa.gr

Abstract. There are two issues that are challenging in the life-cycle of Digital Talking Books (DTB): the automatic labeling of text formatting meta-data in documents and the multimodal representation of the text formatting semantics. We propose an augmented design-for-all approach for both the production and the reading processes of DAISY compliant DTBs. This approach incorporates a methodology for the real-time extraction and the semantic labeling of text formatting meta-data. Furthermore, it includes a unified approach for the multimodal rendering of text formatting, structure and layout meta-data by utilizing a Document-to-Audio platform to render the acoustic modality.

Keywords: e-books, digital talking books, document accessibility, document-to-audio, auditory accessibility of text formatting, DAISY.

1 Introduction

Nowadays, e-books have exceeded printed books by providing browsing, navigation, searching, highlighting and multimedia facilities to the reader. The content is being rendered based on embedded meta-data of structural and layout properties that affect the visual stimuli. The meta-data and the way it is combined and represented in visual modality, has specific meaning for each reader in different documents [1][2] and the way the document is read [3][4]. There are also studies on how the combination of colors on a page and the different type of style on text, affects narrator's emotional state and so the readability of the document [5].

Access to an e-book requires a two-way communication between the user and the e-book. One common format specification is the Digital Talking Book (DTB) [15]. DTB provides multimodal presentation of the content (audio, visual and haptic [16]) and also introduce additional special tagging to e-books in order to facilitate navigation and audio synchronization with the text. DTB can be also used as Digital Audio Books and the content is reproduced (due pre-recorded speech) through common devices like a CD player.

The complexity of accessing e-books in a user-friendly way has led to the proposal of several guidelines and specifications in the field of electronic content accessibility. The NISO/ANSI [17] standard has been proposed for the creation of Digital Talking

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 436–445, 2007.
© Springer-Verlag Berlin Heidelberg 2007

Books. This standard is supported by the DAISY consortium [16] and specifies a faithful access on the representation of a published work to visually impaired and print-disabled readers like to readers of the original printed publication. The Open Document Format (ODF) [18] is an open XML standard to be used for documents containing text, spreadsheets, charts, and graphical elements. W3C also provides recommendations for making web content accessible for the disabled, through Web Content Accessibility Guidelines [19].

In acoustic modality, variations and differences of speech characteristics can have great impact to the perception of the speech. Depending on the listeners and the content that is read, there are different expressive styles of speech [6][7]. Pre-recorded speech and decorative sounds are most commonly used in e-books. This makes the production of an e-book a very time consuming and expensive process but also prevents the dynamic creation of e-books from any available content. Nowadays, synthetic speech has become very natural in quality by the introduction of corpus-based Text-to-Speech (TtS) techniques, but speech synthesis is mono-dimensional and only deals with plain typographic texts. E-books are currently support TtS functionality but the semantics of the original visual format of the book are not successfully transferred to the listener.

Previous works have shown the weaknesses and the errors in the comprehension of the acoustic version of visual-oriented documents when stripping the visual meta-data and synthesizing only the content text [20][21]. Recent works have demonstrated encouraging results in the transformation of the visual stimuli to the acoustic modality. There is a systematic work [22] to develop the Document-to-Audio open platform for vocalizing meta-data along with the content text, producing a two-dimensional audio stream, by: (a) combining alternative text insertion in the document's text stream, (b) altering the prosody, (c) switching voices and (d) inserting other sounds in the waveform stream, according to the class of meta-data provided in the document. A recent study has modeled speech styles when vocalizing tables using synthetic speech [8]. In [9], it is shown how the variations of speech and its characteristics affect the listener's state of emotions (emotional prosody). For the optimization of the acoustic representation, there are efforts that combine speech synthesis with special sounds like earcons [10] and auditory icons [11]. For example, earcons can be used for navigation in a menu [12][13]. Another approach is using 3D Audio for the acoustic representation of the content [14].

In this work we propose an augmented design-for-all approach for both the production and the reading process of DAISY/NISO compliant DTBs. This approach incorporates a methodology for the real-time extraction and the semantic labeling of text formatting meta-data. Furthermore, it includes a unified approach for the multimodal rendering of text formatting, text structure, text layout and non-textual meta-data. Specifically for the acoustic modality, we present a prototype for the auditory presentation of e-book contents.

2 The Proposed Architecture

We introduce an open real-time XML-based architecture allowing a multimodal access to books. The architecture fulfils the needs of two major tasks: a) the

production of compliant DTBs and b) the presentation of their content and its attributes provided by the meta-data. The proposed system is able to process any book, in printed or electronic format.

Figure 1 shows the overall architecture consisting of two parts: the Universal player (U-Player) and the Book Explorer. The later hosts the analysis of the books, as well as the processing of navigation issues. We have followed the client – server model (CS model) in order to serve any small-footprint specifications on the user's device. The U-Player is being installed on user's device (e.g. a personal computer, a mobile device etc). On the other hand, the resource demanding Book Explorer could be hosted on a powerful server machine that performs analysis tasks.

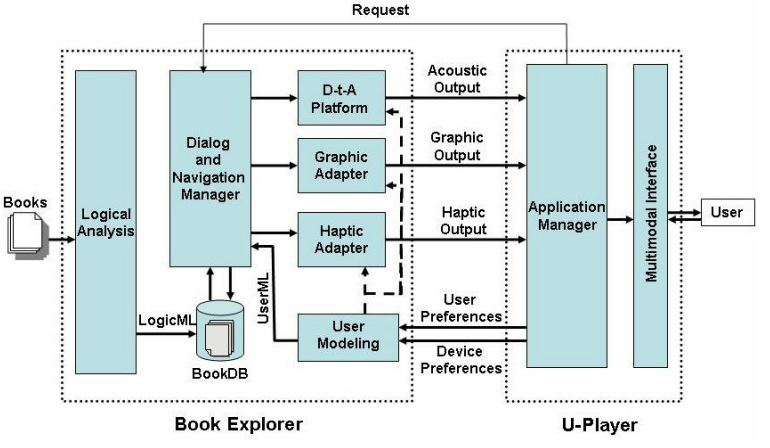


Fig. 1. The overall architecture

The CS model also allows the Server to serve simultaneously multiple clients of different modalities. A user can access the application through internet but another user through a mobile provider, selecting the appropriate gateway for their interface.

2.1 Production Task

The Book Explorer is the module that deals with the production of the e-book and handles the reading and exploration tasks. The Logic Analysis module parses the book (in printed or electronic form) and the produced output (DAISY/NISO standard compliant) is stored in the BooksDB database. The stored content and any additional information (meta-data) is accessed by the Dialog and Navigation Manager. The information retrieval from the BooksDB is done according the demands of the present under execution task. Furthermore, the Dialog and Navigation Manager handles the user and device preferences (UserML) produced by the User Modeling module. Depending on the user preferences it manages the way that content will be handled and delivered in each modality (acoustic, visual, haptic) through the Document-to-Audio platform (or/and the Graphic Adapter and/or the Haptic Adapter) to the Application Manager.

To serve the multimodal accessibility requirements, our methodology first creates a logical layer from the source book, free of presentation details. Then, these logical data can be then transformed to any modality. Though books are aiming at visual modality, this logical layer aims to record the reader's understanding of the book. We will show in detail how the original visual stimuli can affect the acoustic presentation via this layer.

2.2 Presentation Task

The presentation task is carried out at the U-Player component. This takes care for user interaction with the service (and consequently with the content) in multimodal way. The U-Player is divided into two major tasks: input and output.

In output task, the user receives the requested content or other system prompts. The output task can accomplish the type of DTBs with full audio through the Navigation Center (NCC or NCX). The structure is two-dimensional, providing both sequential and hierarchical navigation. In many cases, the structure in this type of DAISY DTB resembles the table of contents of its print source. Some of these productions provide page navigation. But depending on the user's preferences the functionality can change respectively.

In input task the architecture supports multimodal accessibility, depending on the user's needs. The input devices (or applications) vary, from depend on the device Automatic Speech Recognition (A.S.R.), keyboard, mouse or buttons.

The Application Manager handles the user preferences, which are used in Book Explorer or Output tasks. The input can be operation or navigation commands. The application manager creates Requests that contain information about: (a) User preferences, and (b) Navigation or Application Command.

2.3 User Modeling

The User Modeling Module handles the user profile. The profile is a collection of actions and preferences acquired by the interaction with the Application Manager. Such can be the user's needs, the interaction modality, the navigational factors, environmental factors etc.

2.4 Dialog and Navigation Manager

The navigation in DTBs supports several functionalities. As basic navigation is considered the movement through text and sophisticated movement is the hierarchical navigation in book (Using Navigation Control File that describes the hierarchy of the book). The hierarchy is provided separately from the text.

The Dialog - Navigation Manager Module (D.N.-M.) can handle commands from the user for navigation tasks and generally tasks that satisfies user's content and information retrieval. An adaptive dialog system [31] is used that corresponds to the two following functions [32]: a) Supporting System Use, taking over parts of routine tasks, adapting the interface, giving advice of system use, controlling a dialogue.

b) Supporting Information Acquisition, helping users to find information, tailoring information presentation.

This module is agent-based, allowing every module to interact and to be capable of reasoning. This kind of architecture is very flexible and user-centred. The user can have full control of the dialog [33]. Also the proposed architecture is flexible and can support multimodal dialog input and not only spoken dialogs [34], while the navigation features are conforming to the DAISY/NISO standard, supporting Fast Forward and Fast Reverse, Reading at Variable Speed, Notes, Cross Reference Access, Index Navigation, Bookmarks, Highlighting, Excerpt Capability, Searching, Spell-Out Capability, Text Attributes and Punctuation, Tables, Nested Lists, Text Elements, Skipping User-Selected Text Elements, Location Information, Summary and Reporting Information, Science and Mathematics.

3 Logical Document Analysis

The proposed system targets to process all kind of books in printed or electronic format. The Markup Normalization Module converts all non-tagged books as well as tagged books not conforming with the DAISY/NISO format into tagged compliant to DAISY/NISO format. Printed books are scanned and parsed through an OCR system so to be digitized and exported in a tagged format. Documents being already in a tagged format include meta-data about the format and the structure of the text. E-books are also tagged digital books but have specific file format and tagging. All these have to be normalized to the required meta-data and file type (DAISY/NISO standard). This format is described in our methodology as BooksML.

In case the book has the required by our approach semantic annotation (described later), the logic analysis can be by-passed, and the e-book can be stored directly to systems database (BooksDB) (Figure 2).

3.1 Type of Elements

The e-books include meta-data that can be categorized in:

- Text Formatting meta-data
- Text Structure meta-data
- Text Layout meta-data
- Non-textual meta-data (Figures, Drawing, Pictures, Logos etc)

Text Formatting meta-data include the formation elements of the text, typesetting elements and font elements (like bold, italics, font size). Text Structure meta-data specify the attribute of a part of the document (chapter, title, paragraph e.t.c.), while the Text Layout meta-data concerns the formation of the text (like columns, headlines, borders). The text with the formation and structural metadata can be combined with other non-textual metadata, such as figures, drawing e.t.c.

One can notice that there is a relation between these elements. For example, the title (text structure) is in font size 16pt (text formatting) and it is placed in a column (text layout). Also, the subtitle has font size 14pt and it is in a column, but under the

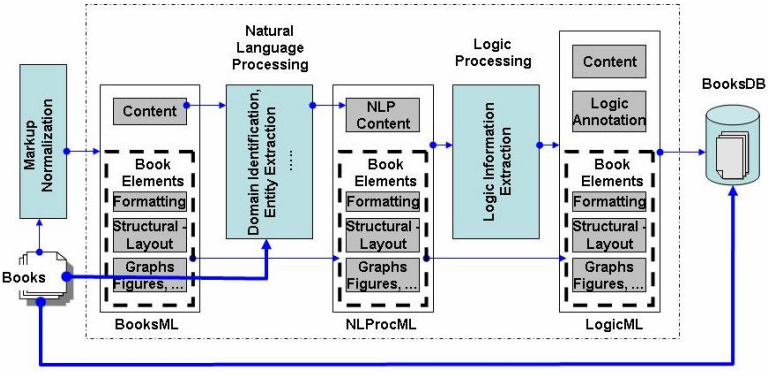


Fig. 2. The Logic Analysis Module. Each sub-system produces an XML file which contains the information that derived by its procedure.

title. We can distinguish that between these elements, the relation is M:N. But there are combinations that can be excluded through a statistical analysis.

The above elements should be included in the tagged book (BookML). The elements that are described in the DTBs standard [28] can be considered as a sub-group of the elements that can be processed by our proposed methodology.

3.2 Semantics of Meta-data

The elements of a document are set by the editor for different purposes. Stricter typographic rules are used in technical documents than in magazines. Not considering the domain of the document, we can see a few examples. A sentence written in bold, size 16pt and on the top of the page is the title of a chapter and gives to the reader the main concept that will be analyzed in the body of the text. A bold, numbered sentence under an image is the caption of the image explaining the image above it.

As we can see, the combination of elements has a semantic mapping. DAISY/NISO standard supports a plethora of semantic meta-data for representation or navigation. Though in acoustic modality these meta-data can be exploited, in this work we introduce additional semantics that are more useful for the acoustic rendering of the e-book. For example, several level of emphasis in synthetic speech are required to produce a more augmented acoustical image of the document compared to the emphasis and strong emphasis levels provided by the DAISY/NISO standard. Other semantic – logic meta-data includes “dialog” (“question” and “answer”), “definition of a word”, “new term” e.t.c.

Logic Information Extraction adds logic meta-data to the e-book using machine learning algorithms (e.g. Bayesian Network) to learn the mapping between visual and logic elements. The statistical analysis can be done through a series of experiments to allocate the mapping, which in some cases is 1:N (e.g. bold can be mapped as emphasis but also as strong emphasis). A questionnaire can be created and given to readers to annotate the logic of the sections in books.

4 Delivering Books into Acoustic Modality

When conveying a book in auditory modality, the functionality is focused on representing plain text stripping all visual, structural and layout information included. So the acoustic presentation lacks of the additional information that is provided. In DTBs for navigation purposes the structural elements are spoken out separately from text.

While the reader uses the navigation features provided in a DTB, it is required these information – metadata conveyed in acoustic modality. The goal is not to separate or remove the metadata from the text itself. Trying to accomplish this goal arises the need of defining which auditory elements these features affect. For example, the title should be spoken out in a different manner than the body of text, changing the appropriate prosodic elements of the synthetic speech or using specific sounds like earcons, so the user to understand which is the attribute of the text that was spoken out without to be defined separately.

In this section we will describe the acoustic mapping of semantic meta-data in books. As we described above, the book is formatted in LogicML. Following previous works on Document-to-Audio conversion [30], the semantics meta-data can be acoustically represented by specific auditory elements like (a) alternative text insertion in the document's text stream, (b) modifications in the prosody, (c) switching voices and (d) inserting other sounds like earcons and auditory icons in the waveform stream, according to the class of meta-data provided in the e-book. The user can be trained to recognize and to combine speech and sounds with specific commands and events.

The book (in LogicML) through the Dialog and Navigation Manager is passed to the Document-to-Audio platform (Fig. 3) [29][30]. It will be mapped in specific acoustic elements, as mentioned above, producing a new annotated document and auditory synthesizer will implement the mapping (the output files can be e.g. Mpeg4, SMIL and audio files).

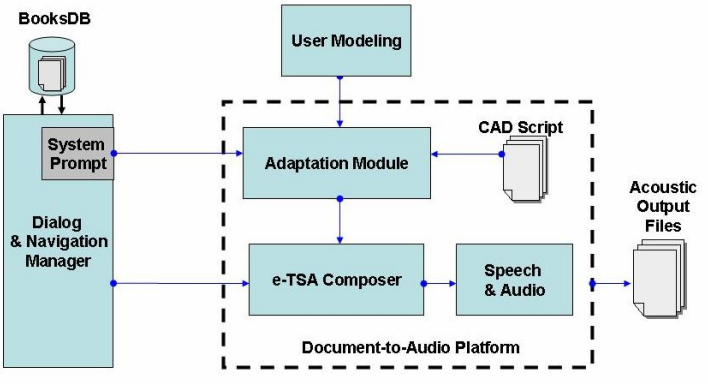


Fig. 3. The Acoustic Mapping module

The mapping will be done according the information provided to DtA platform. The rules that describe the semantic relation of metadata and acoustic representation are defined with the use of Cluster Auditory Definition (CAD) scripts as it is proposed in e-TSA composer platform. Through CAD scripts it is possible to be defined the prosodic behavior and audio insertions, like earcons, in each pair of metadata along with its content text. The DtA platform will give greater priority to user preferences than the default CAD rules. For example, the user might need to hear faster the content of the book but there are elements that should be read slower (defined by CAD). Then the Adaptation Module will give higher priority to the rules provided by User Modeling. Another functionality of Adaptation Module is how to handle System Prompts. The system prompts can be mapped to different elements than these of the content, so to be distinct to the listener.

The rules of mapping can derive either by a "test and error" process or from a series of experiments, like in [8] [26].

5 Conclusions

We have presented an augmented approach for both the production and the reading processes of DAISY/NISO compliant DTBs. Our methodology includes a system architecture for multimodal accessibility of books. In this paper we have emphasised the acoustic modality. Our aim was to propose a methodology for presenting the independency of the logic, hidden in the elements of a book for multimodal accessibility.

Acknowledgements. The work described in this paper has been funded by the European Social Fund and National Resources under the HOMER project of the Competitiveness Programme: PENED, Greek General Secretariat of Research and Technology.

References

1. Küpper, N.: Recording of Visual Reading Activity: Research into Newspaper Reading Behaviour (1989) Available as pdf from http://calendardesign.de/leseforschung/eyetrackstudy.pdf
2. Holmberg, N.: Eye movement patterns and newspaper design factors. An experimental approach. Master Thesis, Lund University Cognitive Science (2004)
3. Holmqvist, K., Holsanova, J., Barthelson, M., Lundqvist, D.: Reading or scanning? A study of newspaper and net paper reading. In: Hyönä, J.R., Deubel, H. (eds.) The mind's eye: cognitive and applied aspects of eye movement research, pp. 657–670. Elsevier, Amsterdam (2003)
4. Holmqvist, K., Wartenberg, C.: The role of local design factors for newspaper reading behaviour – an eye-tracking repspective. Daily newspaper layout-designer's prediction of reader's visual behaviour- a case study Lund University Cognitive Studies, 127. Lund: LUCS (2005)

5. Laarni, J.: Effects of color, font type and font style on user preferences. In: Stephanidis, C. (ed.) Adjunct Proceedings of HCI International 2003, pp. 31–32. Crete University Press, Heraklion (2003)
6. House, D., Bell, L., Gustafson, K., Johansson, L.: Child-directed speech synthesis: Evaluation of prosodic variation for an educational computer program. In: Proc. Eurospeech 1999, Budapest, Hungary (1999)
7. Johnson, W.L., Narayanan, S., Whitney, R., Das, R., Bulut, M., LaBore, C.: Limited domain synthesis of expressive military speech for animated characters. In: IEEE Speech Synthesis Workshop, Santa Monica, USA (2002)
8. Spiliotopoulos, D., Xydas, G., Kouroupetroglou, G., Argyropoulos, V.: Experimentation on Spoken Format of Tables in Auditory User Interfaces. Universal Access in HCI. In: Proceedings of the 3d International Conference on Human-Computer Interaction (HCII-2005), Las Vegas, USA, vol. 8 (July 22-27, 2005)
9. Schröder, M.: Expressing degree of activation in synthetic speech. IEEE Transactions on Audio, Speech and Language Processing 14(4), 1128–1136 (2006)
10. Kramer, G. (ed.): Auditory Display: Sonification, Audification, and Auditory Interfaces. In: Proceedings Santa Fe Institute Studies in the Sci, Addison – Wesley, vol. 18 (1994)
11. Gorny, P.: Typographic semantics of Webpages Accessible for Visual Impaired Users, Mapping Layout and Interaction Objects to an Auditory Interaction Space. In: International Conference on Computer Helping with Special Needs, pp. 17–21 (2000)
12. Brewster, S.A., Raty, V.-P., Kortekangas, A.: Earcons as a Method of Providing Navigational Cues in a Menu Hierarchy. In: Proceedings of HCI'96, Imperial College, London, UK, pp. 167–183. Springer, Heidelberg (1996)
13. Mynatt, E.D.: Designing with auditory icons: how well do we identify auditory cues? In: Plaisant, C. (ed.) Conference Companion on Human Factors in Computing Systems, Boston, Massachusetts, United States, April 24–28, 1994, pp. 269–270. ACM Press, New York (1994)
14. Djennane, S.: 3D-Audio News Presentation Modeling. In: User Interfaces for All, pp. 280–286 (2002)
15. Duarte, C., Carriço., L.: Identifying adaptation dimensions in digital talking books. In: Proceedings of the 9th international conference on Intelligent user interface IUI '04 (2004)
16. Daisy Consortium http://www.daisy.org
17. ANSI/NISO.: Specifications for the digital talking book (2002) http://www.niso.org/standards/resources/Z39-86-2002.html
18. OASIS, Organization for the Advancement of Structured Information Standards http://www.oasis-open.org/home/index.php
19. Web Content Accessibility Guidelines, W3C http://www.w3.org/TR/WAI-WEBCONTENT/
20. Raman, T.V.: An Audio view of (LA)TEX Documents. In: Proceedings of the Annual Meeting. TUGboat, vol 13(3), pp. 65–70 (1992)
21. Hakulinen, J., Turunen, M., Rih, K-J.: The Use of Prosodic Features to Help Users Extract Information from Structured Elements in Spoken Dialogue Systems. In: Proceedings of ESCA Tutorial and Research Workshop on Dialogue and Prosody, pp. 65–70 (1999)
22. Xydas, G., Argyropoulos, V., Karakosta, T., Kouroupetroglou, G.: An Open Platform for Conducting Psycho-Acoustic Experiments in the Auditory Representation of Web Documents. In: Proc. Conference ACOUSTICS (2004)
23. Stephanidis, C., Akoumianakis, D., Sfyrakis, M., Paramythis, A.: Universal accessibility in HCI: Process-oriented design guidelines and tool requirements. In: Proceedings of the 4th ERCIM Workshop on User Interfaces for All, Stockholm, Sweden (1998)

24. EDEaN.: Design for All Network of Excellence (IST-2001-38833) European Design for All e-accessibility Network (EDEaN) (2001) http://www.d4allnet.gr/
25. Europe's Information Society, Design for All http://europa.eu.int//information_society/policy/accessibility/dfa/index_en.htm
26. Xydas, G., Argyropoulos, V., Karakosta, T., Kouroupetroglou, G.: An Experimental Approach in Recognizing Synthesized Auditory Components in a Non-Visual Interaction with Documents. In: Proceedings of the 11th International Conference on Human-Computer Interaction (HCII2005), Las Vegas, Nevada SA, 22-27 July 2005, vol. 3, pp. 411–420 (2005)
27. Kilov, H.: Semantics in open hypermedia systems: Information modeling for document understanding. In: ECHT'94 Workshop on Open Hypermedia Systems Edinburgh, Scotland (1994)
28. Kerscher G.: Theory Behind the DTBook DTD. Daisy Consortium (2001) http://www.daisy.org/ publications/docs/theory_dtbook/theory_dtbook.html
29. Xydas, G., Kouroupetroglou, G.: Text-to-Speech Scripting Interface for Appropriate Vocalisation of e-Texts. In: Proceedings of the 7th European Conference on Speech Communication and Technology (EUROSPEECH 2001) Aalborg, Denmark pp. 2247–2250 (2001)
30. Xydas, G., Kouroupetroglou, G.: Augmented Auditory Representation of e-Texts for Text-to-Speech Systems. In: Matoušek, V., Mautner, P., Mouček, R., Tauser, K. (eds.) TSD 2001. LNCS (LNAI), vol. 2166, pp. 134–141. Springer, Heidelberg (2001)
31. Fink, J., Kobsa, A., Nill, A.: Adaptable and Adaptive Information Provision for All Users, Including Disabled and Elderly People (1999) Available from http://citeseer.ist.psu.edu/fink99adaptable.html
32. Hjalmarsson, A.: Adaptive Spoken Dialog Systems. In: GSLT, Speech Technology 1 Closing Seminar (2005)
33. McTear, M.F.: Spoken dialogue technology: enabling the conversational user interface. ACM Comput. Surv. 34(1), 90–169 (2002)
34. Turunen, M., Hakulinen, J., Räihä, K., Salonen, E., Kainulainen, A., Prusi, P.: An architecture and applications for speech-based accessibility systems. IBM Syst. J. 44(3), 485–504 (2005)

Enhancing the Usability of Real-Time Speech Recognition Captioning Through Personalised Displays and Real-Time Multiple Speaker Editing and Annotation

Mike Wald[1] and Keith Bain[2]

[1] Learning Technologies Group, School of Electronics and Computer Science, University of Southampton, Southampton SO171BJ, United Kingdom
M.Wald@soton.ac.uk
[2] Liberated Learning, Saint Mary's University, Halifax, NS B3H 3C3, Canada
Keithbain@stmarys.ca

Abstract. Text transcriptions of the spoken word can benefit deaf people and also anyone who needs to review what has been said (e.g. at lectures, presentations, meetings etc.) Real time captioning (i.e. creating a live verbatim transcript of what is being spoken) using phonetic keyboards can provide an accurate live transcription for deaf people but is often not available because of the cost and shortage of highly skilled and trained stenographers. This paper describes the development of a system that can provide an automatic text transcription of multiple speakers using speech recognition (SR), with the names of speakers identified in the transcription and corrections of SR errors made in real-time by a human 'editor'.

Keywords: Real time, captioning, speech recognition, editing, multiple speakers, transcription.

1 Introduction

Text transcriptions of the spoken word can benefit deaf people and also anyone who needs to review what has been said (e.g. at lectures, presentations, meetings etc.) Real time captioning (i.e. creating a live verbatim transcript of what is being spoken) using phonetic keyboards can provide a live transcription for deaf people and can cope accurately (e.g. >98%) with people talking at up to 240 words per minute but is often not available because of the cost and shortage of highly skilled and trained stenographers [1] [2]. This paper describes the development of applications that use speech recognition to provide automatic text transcriptions.

2 Visual Indication of Pauses

Standard speech recognition (SR) software (e.g. Dragon, ViaVoice [3]) was found to be unsuitable for live transcription of speech as without the dictation of punctuation it

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 446–452, 2007.
© Springer-Verlag Berlin Heidelberg 2007

produced a continuous unbroken stream of text that was very difficult to read and comprehend. IBM and Liberated Learning (LL) therefore developed ViaScribe [4] [5] as an SR application that automatically formats real-time text captions from live speech with a visual indication of pauses. Detailed feedback from students with a wide range of physical, sensory and cognitive disabilities and interviews with lecturers [6] showed that both students and teachers felt this approach improved teaching and learning as long as the text was reasonably accurate (e.g. >85%).

3 Personalised and Customisable Display

While projecting the text onto a large screen in the classroom has been used successfully in LL classrooms it is clear that in many situations an individual personalised and customisable display (e.g. font size, formatting, colour etc.) would be preferable or essential and so a personalised server and client was developed to enable users to customise their displays on their own networked computer [7].

4 Real-Time Editing

SR accuracy may be reduced where the original speech is not of sufficient volume/quality (e.g. poor microphone position, telephone, internet, television, indistinct speaker) or when the system is not trained (e.g. multiple speakers, meetings, panels, audience questions). An experienced trained 're-voicer' repeating what has been said can sometimes improve SR readability in these situations by correcting ASR errors if the accuracy is high and the speaking rate low and summarising what is being said if the speaking rates are fast [8] [9]. Summarisation however requires the re-voicer to actually understand and 'interpret' what is being said and therefore to have a good knowledge of the subject.

To improve accuracy of verbatim captions created directly from the voice of the original speaker the application RealTimeEdit (RTE) was developed to enable corrections to ASR captions to be made in real-time [10]. One editor can find and correct errors or the task of finding and correcting errors can be shared between two editors, one using the mouse and the other the keyboard. It is also possible to use multiple editors sequentially to allow a 2nd operator to correct errors that a 1st operator didn't have time to correct. The editor can also annotate where required (e.g. describe sounds <<LAUGHING>> or identify mumbled and clearly incorrectly recognised words that they cannot identify as <<INAUDIBLE>>). In this way a real-time editor can be used in situations where high accuracy captions are required and a real-time stenographer is not available. Up to eleven corrections per minute were achieved by untrained users of an initial prototype of RTE and a theoretical analysis suggested experienced touch typists could be trained to achieve over 15 corrections per minute. Analysis of an ASR transcript with a 22% error rate also suggested that correction of less than 20% of the 'critical' errors may be required to understanding

the meaning of all the captions [11]. Somebody talking at 150 words per minute with a 22% error rate produces an average of 33 errors per minute and if correction of only 20% of these errors were 'critical' to understanding then the editor would have to correct on average only about 7 errors per minute. This would suggest that even if 100% accuracy was not achievable, 100% understanding might be. Judging which words were critical to understanding might be an easier task than the summarisation task faced by the re-voicer.

5 Multiple Speaker Transcriptions

In situations where there is more than one person speaking, using multiple instances of ViaScribe creates captions in multiple windows making it difficult to follow the sequence of the utterances. To produce a transcript of the session with speakers identified, the application RealTimeMerge (RTM) was developed to add the speaker's name to the text captions and merge the streams from the instances of ViaScribe. Each speaker and instance of ViaScribe can have a separate editor and the edited outputs merged or the unedited outputs of ViaScribe can be edited. The combination of ViaScribe, ViaScribe server, PDC, RTE, and RTM enables a very flexible

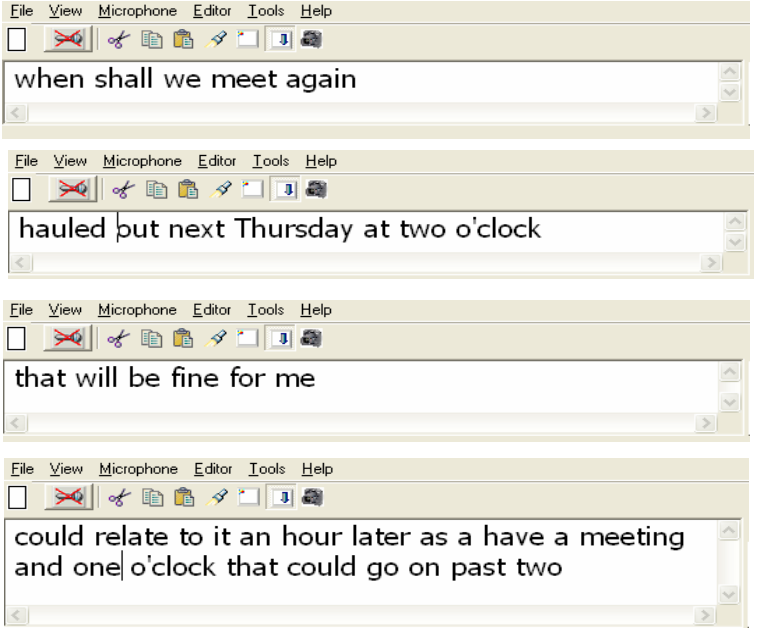


Fig. 1. Four Instances of ViaScribe showing the ASR text captions with errors for four separate speakers

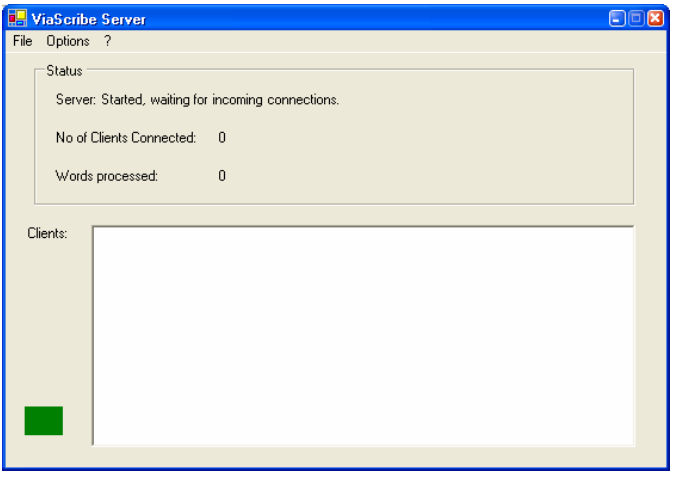


Fig. 2. ViaScribe Server started

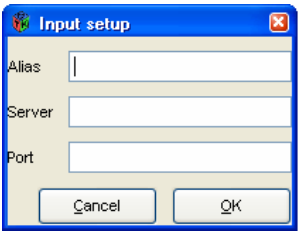


Fig. 3. RealTimeMerge Input setup window

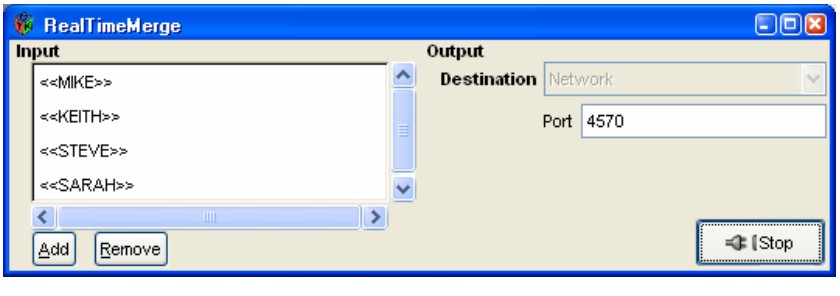


Fig. 4. The four speakers' names and Server IP addresses and Port numbers have been added to RealTimeMerge

approach to be adopted that can provide solutions to many requirements. Figures 1-7 show how the recognised text from four speakers using four instances of Viascribe can be output via the server and merged with the speaker's names added and then edited for errors before the corrected transcript is displayed on one or more clients.

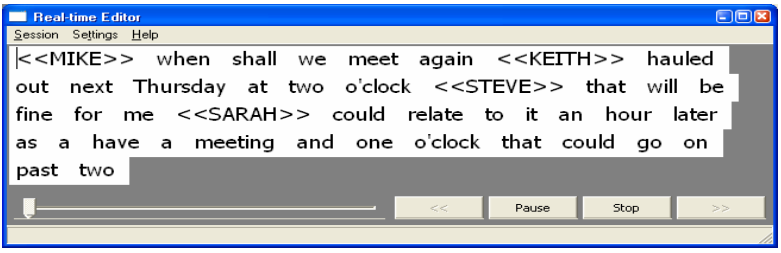


Fig. 5. RealTimeEdit displaying the merged captions and names of the four speakers' output by RealTimeMerge

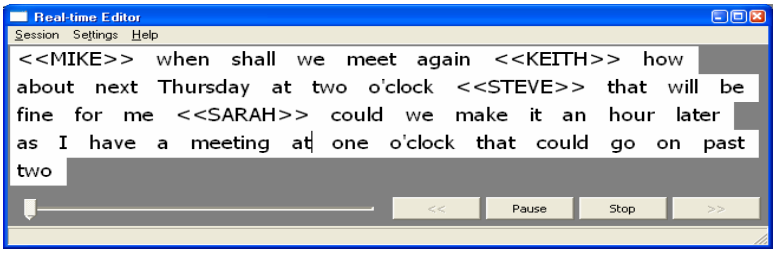


Fig. 6. RealTimeEdit displaying the captions after correction

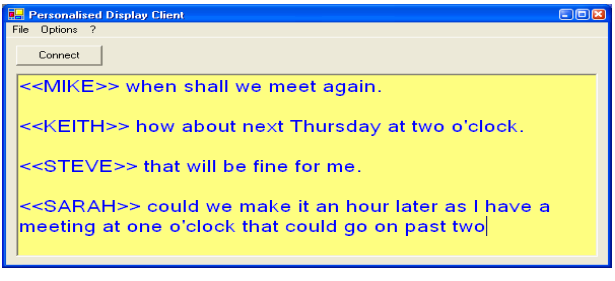


Fig. 7. Personalised Display Client displaying the corrected captions sent by RealTimeEdit

Figure 8 shows some of the many possible configurations while the following bullet points provide some further details:

- The RTEs and RTM and PDC can be on the same computer as the Server or each other OR on a different computer that is on the same network but the Server must be on the same computer as ViaScribe;
- One Client (A) can be connected to the Server to show the unedited ViaScribe display without a delay while another Client (B) can be connected to RTE to show the edited display with an editing delay. Another client (e.g. C) can be connected to RTM to show the unedited merged displays of multiple speakers while another client (e.g. D) can be connected to RTE to show the edited merged display with an editing delay;

- RTEs can be 'daisy chained' to allow a 2nd operator to correct errors that a 1st operator didn't have time to correct). For example a Client (E) (e.g. for a user who required greater accuracy because they were profoundly deaf) could be connected to RTE (ii) that takes as its input the output of RTE (i);
- Each RTE can have either one 'operator' to find and correct errors OR the task of finding and correcting errors for each RTE can be shared between two operators, one using the mouse and the other the keyboard;
- The unedited outputs of ViaScribe can be merged and then edited as shown in figure 8 or if preferred, each instance of ViaScribe can have its output edited using RTE and then all the edited RTE outputs can be merged;

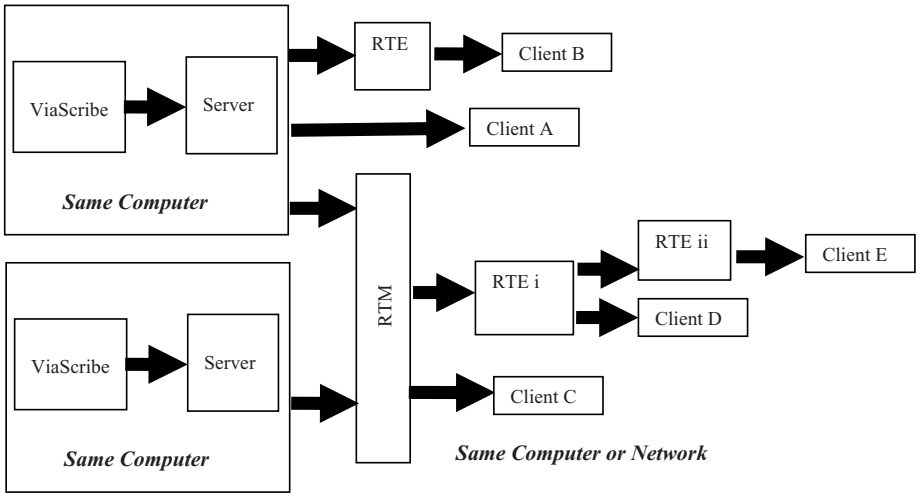


Fig. 8. Diagram showing how ViaScribe, Server, RTE, RTM and PDC can be configured

6 Further Work

Trials of the system in a variety of settings are being conducted to investigate in practice the effect of error rates, number of speakers, editing operator skill requirements etc.

7 Conclusion

A range of applications have been developed to enhance the useability, readability and accessibility of SR for real-time speech to text transcription. The combinations of the Server, PDC, RTE and RTM provide a very flexible approach for many different requirements and trials of the system are being conducted to determine how well the system can cope with the complexities of real environments.

References

1. Wald, M.: An exploration of the potential of Automatic Speech Recognition to assist and enable receptive communication in higher education. ALT-J, Research in Learning Technology 14(1), 9–20 (2006)
2. Wald, M., Bain, K.: Using Automatic Speech Recognition to Assist Communication and Learning. In: Proceedings of HCI International 2005: 11th International Conference on Human-Computer Interaction, Las Vegas USA. vol. 8 (2005)
3. Nuance (2006) (Retrieved February 7, 2007) from http://www.nuance.co.uk/
4. Bain, K., Basson, S., Wald, M.: Speech recognition in university classrooms. In: Proceedings of the Fifth International ACM SIGCAPH Conference on Assistive Technologies, pp. 192–196. ACM Press, New York (2002)
5. IBM (2005) (Retrieved February 7, 2007) from http://www-306.ibm.com/able/solution_offerings/ ViaScribe.html
6. Leitch, D., MacMillan, T.: Liberated Learning Initiative Innovative Technology and Inclusion: Current Issues and Future Directions for Liberated Learning Research. Saint Mary's University, Nova Scotia. (2003) (Retrieved February 7, 2007) from http://www.liberatedlearning.com/
7. Wald, M.: Personalised Displays. In: Speech Technologies: Captioning, Transcription and Beyond IBM T.J. Watson Research Center New York USA (2005) (Retrieved February 7, 2007) from http://www.nynj.avios.org/Proceedings.htm
8. Lambourne, A., Hewitt, J., Lyon, C., Warren, S.: Speech-Based Real-Time Subtitling Service. International Journal of Speech Technology 7, 269–279 (2004)
9. Francis, P.M., Stinson, M.: The C-Print Speech-to-Text System for Communication Access and Learning. In: Proceedings of CSUN Conference Technology and Persons with Disabilities. California State University Northridge (2003)
10. Wald, M.: Creating Accessible Educational Multimedia through Editing Automatic Speech Recognition Captioning in Real Time. International Journal of Interactive Technology and Smart Education: Smarter Use. of Technology in Education 3(2), 131–142 (2006)
11. Wald, M.: Research and development of client-server personal display of speech recognition generated text, real time editing and annotation systems: Speech Technologies-Accessibility Inroads: A special symposium on accessibility and speech recognition technology. IBM Hursley Research Park (2006) (Retrieved February 7, 2007) from http://www.liberatedlearning.com/news/proceedings.html

A Study on the Straight-Line Drawing Tasks for the Non-sighted People

Chih-Fu Wu[1], Cherng-Yee Leung[1], Yung-Hsiang Tu[1], Kuang-Chieh Hsu[2], and Huei-Tang Lin[2]

[1] Faculty, Department of Industrial Design, Tatung University, 40, Chungshan N. Rd., Sec. 3, Taipei, Taiwan
[2] Graduate student, Department of Industrial Design, Tatung University, 40, Chungshan N. Rd., Sec. 3, Taipei, Taiwan
{Wcf,Leung,Tys}@ttu.edu.tw, {G9504018,E9504010}@ms2.ttu.edu.tw

Abstract. A new ruler employs the finger touch to be the main pointing sensor with two locators at the two ends, and a unique pen in a smooth track to draw a line. The tasks of making straight line with two difference rulers (Ruler S: ruler used at school; Ruler C: new touch ruler) between non-sighted people (blindfolds and blinds) were observed: (1) free-line; (2) constrained-line; (3) extension-line; (4) closing-line; and (5) conjoining-line. Stages of those tasks were measured: searching, positioning, plotting, and checking, with reacting time and the deviation. The Paired *t*-test shown Ruler S had a larger reacting time for blinds at task (4), (5). MANOVA shown the blinds had larger reacting time at task (3), while (5) on deviation; Ruler C had a larger error at start- point of task (3) and a larger length error of task (3); while Ruler S had a larger error at end-point of task (4). Conclude that some revised design of the ruler really affected the performance of the straight-line tasks, while some did not.

Keywords: Industrial design, Straight-line drawing, non-sighted people, ruler.

1 Introduction

The blinds have to spend more years of learning on the geometric drawing, owing to the lack of visual contacts on the pictures of the textbooks and on the teaching aids [6] [10] [13]. Researchers, [11] [14], had studied on the issues of learning the geometric drawings for blinds for few years. It showed that the blinds usually have to touch all of the pictures to understand the overall shape of pictures, then gathering the oral words to realize the knowledge behind the pictures. Many works, [5] [12], had contributed on the differences of perception of tactual pictures of 2D and 3D, and some others dedicated on the recognition of different views of pictures, such as side-view and perspective view, and some works approached the sizes, scales, and complexity of the drawings showed to the blinds [1] [3]. Those studies focused on the overall recognition ability of participants about the pictures they supported or the pictures the participants drew, in some way, their objectives were to find out the ability of picture discrimination between the subjects. The study here was focused on

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 453–461, 2007.
© Springer-Verlag Berlin Heidelberg 2007

the precision issue to draw a picture at a right size, differed from works before, and proposed a method of comparison about the outcome of non-sighted participants.

Pictures drawn by the blinds usually mixed with some different parts of an object come up with their memory, sometimes with different views of the object they percept, so studies involved pictures drawn by blinds usually had to have the subject to explain about the picture they drawn, not to mention of gaining a exact size of picture the blinds attempt to draw. On the observations of making pictures by blinds, the line thrashed and the endpoints opened usually [7] [4]. It might be caused by lacking of good tools to aid them to predict the direction of their line, and this raised the idea of this work. The purpose of this article was to compare two straight line rulers, with different ways of approaching, on aiding the non-sighted to draw different types of lines, and reported the phenomena found the experiment.

Table 1. The types of straight line and the stages of making them were studied, and the reacting time, shown as solid circles, was recorded in this work

Task	Stage							
	searching		positioning		plotting		checking	
	Start	End	Start	End	Start	End	Start	End
Free-line								
Constrained-line								
Extension-line								
Closing-line								
Conjoining-line								

2 Methods

It seemed to be a simple task to draw a straight line on the swell paper for a sighted people, but the similar task for a blind to draw a precise line will be a rather difficult job. A five-task of experiment was conducted, and each of the tasks was decomposed into four stages, Table 1. Considering the task to draw a line on a paper, there will be some possible types of doing this. Firstly, a *free-line* with no constrain will be the first place to do, which means no limits about its length and direction. Secondly, a *constrained-line* with a specific length is another type that usually happened for the sighted people, and which is still hard to the blind people for they have to search for the start position and aim the pen on the position to draw the line, and to stop at the exactly point of the end position, with no visual contact onto the start and end point. Thirdly, to draw an *extension-line* that is from a pointed start point to a free position. Fourthly, to draw a *closing-line* that is from a free start point to a specific end position. And finally, to draw a *con-joining line* that is to join two lines with a new line. Those 5 types of line were the major tasks of this experiment, and the participants had to draw all of the lines within this approach with random assigned order. Also when carefully analyzed the process of making a line on the paper, subjects usually doing it through four consecutive stages, *searching, positioning, plotting*, and *checking*. In this study, the time spent in the 4 stages was measured to see the differences between the participants and between the rulers.

Thus, all of the participants were to go through all of the 5 straight-line drawing tasks and to be measured their action time about the 4 possible stages happened in the tasks as the time variables. To know the precision of the lines drawn by the subjects, the result swell papers were to be measured the deviation as the accuracy variable.

2.1 Participants

Within the 24 participants, *Table 2*, in the experiment, 12 were blindfolded recruited from Tatung University, and 12 were blind. Seven of the 12 blinds were males volunteered from the Taipei Association of visual Impairment, whose age were 25 to 55 years old (mean is 37.2), and 5 females whose age were 30 to 42 years old (mean is 37.6). The blindfold group was half males, whose age were 21 to 26 years old (mean is 23.6), and half females, whose age were 20 to 23 years old (mean is 21.8).

Table 2. The 24 subjects recruited in this straight-line experiment

Description	Subject			
	Blinds		Blindfolds	
	female	male	female	male
Amount	5	7	6	6
Age (range)	30~42	25~55	20~23	21~26
Age (mean)	37.6	37.2	21.8	23.6
Blind since	CB 2 LB 3	CB 3 LB 4	(sighted)	(sighted)

2.2 Materials

The two rulers (Fig. 1), *Ruler S* and *Ruler C*, were totally different in shape and manipulation. Ruler S had been used at school for visual impairment now, which had the shape similar with the normal ruler for the sighted, only with Braille and dots unit (1 inch/unit) printed upon the surface and an undercut at each unit on the edge of the ruler for the feedback of plotting. Ruler C was a new design concept created by the authors, with the help of CAID Lab, Tatung University, that proposed a new way to draw a line with 2 locators, as point locators and interrupters of the pen (a unique pen placed between the locators), which were to be found at the two end of the ruler, and with thin tactual rod pointed the unit (1 inch/unit).

The key idea of this ruler C is, to draw a line, firstly to search and position the start point and end point by using the finger touch and move the locator to the proper point, then push the pen freely to plot the line. Hopefully in this manner of operation, the blinds could get an exact straight line they wish to draw.

The ruler was situated on a table (29cm x 29cm) covered with a rubber sheet; a swell paper was affixed upon the rubber cover and changed by the experimenter as those different tasks proceeded. The swell paper was suitable for the study for it allowed smooth lines to be produced with no dots and the lines produced could be percept well by hands [2] [8] [9]. In this study, swell paper prepared for those 5 straight-line tasks were: (1) free-line task: a blank paper, the subject was asked to

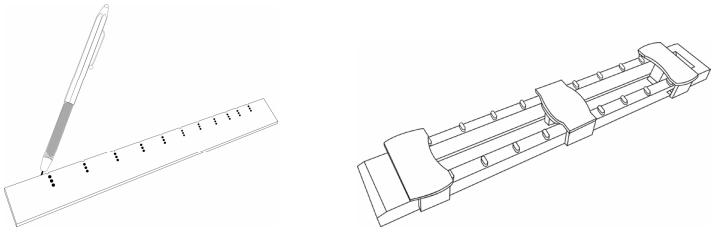


Fig. 1. The straight-line rulers studied in this article, Ruler S, the left one, used at the school of visual impairment, which was made of aluminum with undercuts on the unit position and Braille marks raised on the surface; Ruler C, the right one, a new concept created for the straight-line task, which had 2 indicators at the two ends and a special pen could smoothly move back and forth in between

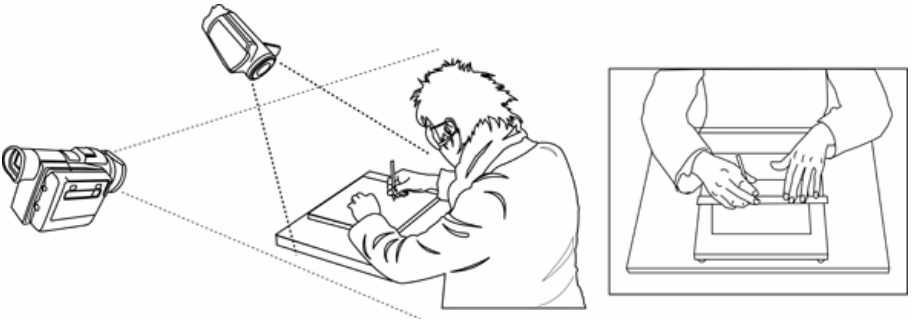


Fig. 2. The experiment was conducted in the study room in CAID Lab., Tatung University with 2 video cameras and 3 experimenters. The right front camera focused on the movement of the major operating hand for the analyzer to determine the reacting time carefully later on.

draw a free length of straight line; (2) constrained-line: another blank paper, the subject was asked to draw a 3 units long of straight line; (3) extension-line: a paper with a dot in the left side, and the subject was about to draw a line from the dot (start-point) to a 3 units away to the right; (4) closing-line: a paper with a dot in the right side, subject was to draw a line from the left side 3 units away the dot to the specific dot (end-point); (5) conjoining-line: 2 dots at the opposite side of the paper with a 5 units distance, the subject was to link the 2 dots with a straight line.

Two video cameras were positioned in the study room (Fig. 2); one was located on the left front side to see the overall action of the participant; the other located on the front top of the subject to observe the hand moving. Three experimenters stood aside in the study room, the first one operated the facade camera to catch the hand movement of the subject in detail; the second one had to change the swell paper before each task proceeded; and the third one read the instruction to the subject and gave the help when subjects needed. The subjects could ask any question when they felt needed in any tasks and the third experimenter would not give any hints unless the subject asked.

2.3 Procedures

The free-line task was to verify whether the subject could actually understand the operations of the rulers, so the data excluded. The other four tasks were randomly assigned to the subjects in order to prevent the possible bias among tasks. The subject was blindfold first, and touched the ruler followed the introduction of its function and operation way by the experimenter. Then the researcher setup the training table for the subject to test the function, and this training phase ended when the subject responded well understood about this ruler.

From this on, the subject was to follow the instruction to use the 1st ruler to draw 5 types of line, and then use the 2nd ruler to draw other 5 types of line in a randomly assigned order. After all the 10 tasks being finished, an interview followed to gather her (his) opinions about the two rulers.

The reacting time of each stages (searching, positioning, plotting, checking) of those tasks (constrained-line, extension-line, closing-line, conjoining-line) were recorded by the video cameras. Because those stages were happened at different tasks, the incision method needed to be standardized, e.g. the time period from the command of beginning till the subject find out the point on the swell paper was recorded to be the searching stage, the time period from the subject moved the ruler to match the point found till the pen picked up was recorded to be the positioning stage, etc., and the questioning time or inessential action time were omitted. Also the deviations of the line drawn were measured included the position error at the start-point and the error at the end-point, also the length of the line drawn.

3 Results

The aims of this study were to evaluate the performance of the 2 straight-line rulers, this involved 2 subject groups, blinds and blindfolds, with 4 straight-line tasks.

The overall means and S.D. were shown in Table 3, and Table 4. The task (2), constrained-line task, had the lowest mean reacting time (38.1 sec.), while the task (5), conjoining-line, had the highest mean (53.8 sec.); the line drawn in the task (2) had the lowest deviation (but the S.D. was very high, 12.4 mm), the task (4) had the largest deviation (interestingly the value shown, -1.7 mm, shorter than it should be); and finally the deviation about the endpoints revealed task (3) and task (5) kept almost the same deviation, but at opposite direction, at the start-point, -0.88 mm and 0.83 mm, respectively, while the task (4) performed better than task (5) at the end-point of the line (1.2 mm compared with -0.6 mm).

Table 3. The overall means and S.D. of reacting time

	T2_ALL	T3_ALL	T4_ALL	T5_ALL
n (people)	48	48	48	48
Mean (sec.)	38.12	51.47	49.41	53.77
SD (sec.)	26.24	38.53	39.43	42.17

Table 4. The overall means and S.D. of deviation

	T2_ERR_L	T3_ERR_L	T4_ERR_L	T5_ERR_L	T3_ERR_S	T5_ERR_S	T4_ERR_E	T5_ERR_E
n (people)	48	48	48	48	48	48	48	48
Mean (mm)	0.00	1.09	-1.68	1.57	-088	0.83	-056	1.29
SD (mm)	12.36	12.29	10.82	7.73	1.77	5.14	8.00	6.12

The analysis was separated into two parts. Firstly, the paired *t*-test was to find out which ruler had better to the subject groups by looking at the reacting time. Secondly, between the tasks, the MANOVA was applied to compare the effect of subjects and rulers at the 4 tasks by looking at the total reacting time and deviation of each task.

3.1 Between Rulers

Of the Blind group, it showed that significant differences shown between the rulers at the task (4), t (11) = 2.30, p=.04, and task (5), t (11) = 2.24, p=.04. Which meant Ruler S > Ruler C at the reacting time of task (4), $M_{Ruler\ S} = 66.6$; $M_{Ruler\ C} = 45.9$; also task (5), $M_{Ruler\ S} = 71.4$; $M_{Ruler\ C} = 45.0$. Although other tasks, (2), (3), did not show clear evidence of improvement, the reacting time of the Ruler C were also shown lower than the Ruler S about 10 sec. at least. This gave a confidence to the design team members of the new ruler (Ruler C) that some part of the new design did improve the straight-line task.

Of the Blindfolds group, however, one part of the data in this group had some hint of defeat about the new design ruler, which was at the task (2) the Ruler C, 44.4 sec., had a larger reacting time than the Ruler S, 29.1 sec. The task (2) was to draw a 3 unit long of line, and the reacting time was optimistically at the blinds group, Ruler S > Ruler C, $M_{Ruler\ S}$ (45.5) > $M_{Ruler\ C}$ (33.5), this indicated that there were still some parts of the new design needed to be adjusted.

3.2 Between Tasks

The MANOVA was employed to analyze the multiple dependent variables measured in this work, included were reacting time of task (2), (3), (4), (5), and deviation of length of task (2), (3), (4), (5), and the deviation of point about start-point of task (3), (5), and about the end-point of task (4), (5), independently within the effects of 2 factors, Subject / Ruler, with 2 treatments in each of them.

The Wilks' Lambda was not significant at Subject, F (12, 33) = 1.07, p= .42, or Ruler, F (12, 33) = 1.70, p= .11, nor the (Subject x Ruler), F (12, 33) = 0.77, p= .67, indicated the interaction between the two factors was not strong enough. The Levene test shown significant differences only at the T3_ERR_L, p=.013, the deviation of length at task (3), and the T3_ERR_S, p=.008, the deviation of start-point at task (3).

At the paired-comparisons between subjects with all the rulers, the T3_ALL, reacting time of task (3), had significant effects of M_{blinds} (63.59) > $M_{blindfolds}$ (39.35), at F (1, 44) = 5.03, p=.03, and also the T5_ERR_E, the deviation of end-point at task (5), had significant effects of M_{blinds} (3.25) > $M_{blindfolds}$ (-0.68), F (1, 44) = 5.27, p=.03. This suggested that the job of task (3) was harder for blinds or somehow there

were some difficulties for them to do it. And the T5_ERR_E measured the deviation of the end-point of the conjoining-line task it seemed that the blindfolds were much more cautious when the pen near to the end-point, so the deviation was less than the blinds.

With all of the subjects to see the effects of the two rulers on those tasks, it was found that Ruler S performed well than Ruler C at T3_ERR_L of $M_{Ruler\ C}$ (4.83) > $M_{Ruler\ S}$ (-2.64), $F(1, 44) = 4.75$, p=.04, also at the T3_ERR_S of $M_{Ruler\ C}$ (-1.69) > $M_{Ruler\ S}$ (-0.07), $F(1, 44) = 11.88$, p=.00, this indicated that Ruler C was not at a good position to help the blinds at the task (3), which was to draw a line from a specific start-point. The other significant differences at this part was the T4_ERR_E of $M_{Ruler\ S}$ (-2.89) > $M_{Ruler\ C}$ (1.77), $F(1, 44) = 4.25$, p=.05, this seemed to be a phenomena that the subjects were afraid to over draw the line asked in length, so at Ruler S, they lifted the pen much earlier than it should be and caused the short in length of this line, and much more to say was the indicator of Ruler C seemed to be a little too loose in holding up the pen that led the pen over drew a little on the swell paper.

4 Conclusions

The 2 rulers were all designed for the tasks of straight-line, especially for visually impaired people, but in different ways. Ruler S showed excellent, simple, of characteristics at plotting, changing the position, and response of units (with the undercuts). While the Ruler C focused on the precision of line and exact endpoints at its original idea. This article was an attempt of reasoning the advantages and defeats of the new concept, Rule C, and by this study this design was encouraged and looking forwards for the next more detailed and useful one yet to come.

4.1 To Draw a Line in a New Way

In general, the new tool opened a different idea to draw a line, i.e. to decide the two end-points first, then bringing two indicators over there to stop the pen, and finally moved the pen freely back and forth and done. The experiment verified the part mentioned that the lines drawn could be straight for the blinds, and with the help of the two indicators the precision should be good enough than the other one.

To draw a line in this case, the 5 tasks were to test overall performance of those rulers, and to find out the directions for the design to be revised. To summarize this work, the next few points have to point out:

(1) As the means shown, the reacting time were Ruler S > Ruler C, in all of the tasks and in all of the subjects, except for the task (2) of blindfolds group, this was a message for our team to follow up with this development for better results.

(2) As the detailed data revealed, at the task (3), to draw a line from the point asked, the new ruler C achieved the lower reacting time but raised at the length error, also the start-point error, whether this was caused by the actions differences or other reasons should be further observed.

(3) As those deviation variables hit, it could not be told between the subjects about the deviation, length drawn and positions, happened at different tasks, except the end-point of task (5) for blindfolds performed well better.

4.2 Implication for Future Work

Following the experiment a face to face interview was performed with the participant, in general, they were interested in the study with the new ruler. But, still they gave the design team many opinions could be launched into the next design, e. g., the indicators were a little bit larger than needed; the groove for the index finger to touch should be wider in case of different tactual surfaces and situations.

More subjects to test the new design will be necessary for gathering more detailed information. It will be a consideration of taking more factors into the study in the future, include wide range of age, profession, and educational level, especially bringing the new ruler to the school of visual impairment for practical evaluation.

Acknowledgments. The study was supported by the National Science Council, NSC 952221E036002 and Tatung University, B94D09024, Taiwan. The authors would like to thank the subjects who volunteered in this evaluation work. And the special thanks go to those research assistants at the CAID Lab. of ID department in the Tatung University.

References

1. Dulin, D., Hatwell, Y.: The Effects of Visual Experience and Training in Raised-Line Materials on the Mental Spatial Imagery of Blind Persons. Journal of Visual Impairment & Blindness, pp. 414–424 (2006)
2. Heller, M.A., Brackett, D.D., Scroggs, E.: Tangible picture matching by people who are visually impaired. Journal of Visual Impairment & Blindness, pp. 349–353 (2002)
3. Heller, M.A.: Picture and pattern perception in the sighted and the blind. The advantage of the late blind. Perception 18, 379–389 (1989)
4. Kamel, H.M., Landay, J.A.: The use of labeling to communicate detailed graphics in a non-visual environment. In: Extended Abstracts of the ACM Conference on Human Factors in Computing, pp. 243–244 (2001)
5. Kawai, Y., Tomit, F.: A support system for the visually disabled to recognize 3D objects. In: Proceedings of Second International ACM Conference on Assistive Technologies, pp. 45–50 (1996)
6. Kennedy, J.M.: Drawing and the blind. Yale University Press, New Havent, C. T (1993)
7. Kurze, M.T.: A computer-based tactile drawing tool for blind people. In: Proceedings of Second International ACM Conference on Assistive Technologies, pp. 261–266 (1996)
8. Lotzsch, J.: Computer-Aided access to tactile graphics for the blind. In: Zagler, W., Busby, G., Wagner, R.R. (eds.) ICCHP 1994. LNCS, vol. 860, pp. 575–581. Springer, Heidelberg (1994)
9. Nomad, P. D.: In audio-tactile tool for the acquisition, use and management of spatially distributed information by visually impaired people. In: Second International Symposium on Maps and Graphics for visually Handicapped People (1988)

10. Schiff, W., Foulke, E.: Tactual perception: a sourcebook. Cambridge University Press, Cambridge (1982)
11. Tu, Y.H., Wu, C.F., Yeh, I.T.: The effects of object name perception and object represented drawing on the recognition of tactile pictures. In: 2005 International Design Conference. p. 37 (2005)
12. Tu, Y.H., Wu, C.F., Yeh, I.T.: Evaluation of Recognizing Tactile Pictures in Different Size Display in Sighted and Blind People. In: The 6th Asian Design International Conference (2002)
13. Warren, D.H.: An individual differences approach. Blindness and children (1994)
14. Wu, C.F., Tu, Y.H., Yeh, I.T.: A Study of Elements Influencing the Visually Impaired on Recognition of Haptic Pictures. In: The 11th International Conference Human-Computer Interaction p. 1273 (2005)

Performance of Different Pointing Devices on Children with Cerebral Palsy

Ting-Fang Wu[1] and Ming-Chung Chen[2]

[1] Graduate Institute of Rehabilitation Counseling, National Taiwan Normal University,
162, Ho-Ping E. Rd., Sec.1 Taipei, Taiwan, R.O.C.
tfwu@ntnu.edu.tw
[2] Department of Special Education, National Chiayi University,
300 University Rd., Chiayi, Taiwan, R.O.C.
mtchen@mail.ncyu.edu.tw

Abstract. With the popularization of graphical user interfaces (GUIs), pointing devices have become standard equipment of most computer systems. However, for people with motor impairments, accurate cursor control is often difficult or impossible to achieve. The focus of this study is to empirically compare effectiveness of three different pointing devices on children with cerebral palsy in order to select the most efficient device for each individual. Computerized Assessment Tool was developed to assess clients' cursor control performance and to provide proper recommendations of pointing devices based on the results of assessments. Performance of pointing, clicking, and dragging of using different devices was measured. The results of this study indicated that Joystick seems to be the most effective device for children with cerebral palsy to perform cursor control tasks. However, factors such as each participant's experience and preference may be related. The instrument use in this study, CAT, seems to be a convenient tool for rehabilitation therapists to assess clients' cursor control performance. Replication of this study using a larger sample size, or recruiting clients with different diagnosis and various ages should be developed.

Keywords: Pointing Devices, Children with Cerebral Palsy, Computerized Assessment Tool, Mouse, Trackball, Joystick, Cursor Control Movement.

1 Introduction

With the popularization of graphical user interfaces (GUIs), pointing devices have become standard equipment of most computer systems. However, for people with motor impairments, accurate cursor control is often difficult or impossible to achieve [1]. Only few research focus on the problems that users with motor difficulties experience with standard mouse. Hwang, Keates, Langdon, and Clarkson (2005) compared mouse movements of six motion-impaired users and three able-bodies users and reported that some motion-impaired users pause more often and for longer than able-bodied, and require up to five times more submovements to complete the same task, and exhibit a greater decline in accuracy with increasing speed than able-bodies

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 462–469, 2007.
© Springer-Verlag Berlin Heidelberg 2007

users [2]. Trewin & Pain (1999) reported the mouse errors of twenty persons with motor disabilities and indicated that the majority of the subjects observed found the difficulty of using mouse and many generally try to avoid mouse operations as much as possible. Person with disabilities reflected in high error rates for positioning the mouse in clicking and dragging operations, and long times spent positioning and dragging relative to the compared group [1].

In order to develop more effective means of computer access, rehabilitation professionals are engaged in selecting appropriate interacting devices and establishing barrier-free computer operation environment for persons with physical disabilities [3]. The focus of this study is to empirically compare effectiveness of three different pointing devices on children with cerebral palsy. Performance of pointing, clicking, and dragging of using different devices was measured.

2 Method

2.1 Participants

Six children with cerebral palsy participated in this experiment with their parents' consents. Their ages ranged from 6 years 10 months to 10 years 2 months with an average of 8 years 5 months. All participants were diagnosed with cerebral palsy. Detailed personal characteristics are described in Table 1.

Table 1. Participant characteristics

Participate	Diagnosis	Sex	Age
A	Athetoid Cerebral Palsy	Female	10 years 2 months
B	Mixed type Cerebral Palsy	Female	7 years 9 months
C	Athetoid Cerebral Palsy	Female	6 years 10 months
D	Spastic Quadriplegia Cerebral Palsy	Male	8 years 2 months
E	Athetoid Cerebral Palsy	Male	9 years
F	Spastic Diplegia Cerebral Palsy	Female	8 years 9 months

2.2 Instruments

The instruments used in this study include a self designed computer software program "Computerized Assessment Tool "(CAT) and three different pointing devices.

Computerized Assessment Tool (CAT). CAT was developed to assess clients' cursor control performance and to provide proper recommendations of pointing devices based on the results of assessments [4] [5]. CAT was developed by C++ Builder, and it consists of an assessment program and a database to record client's information and assessment process. CAT provided a convenient tool for rehabilitation therapists to assess clients' cursor control performance by asking clients to perform clicking, pointing, and dragging with various icon sizes, different cursor moving distances and directions. The figure 1 shows the entry screen layout of the CAT system. The evaluators can set up the evaluation tasks before the assessments

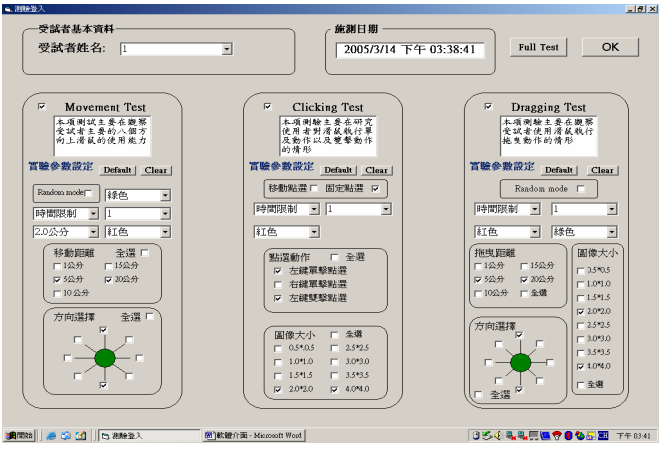


Fig. 1. The screen layout of the CAT system

based on the individual's needs. The Access database collected the client's accurate rate, movement time, cursor movement trajectory, and the speed of cursor movements.

Pointing Devices. Three pointing devices: a Logitech Wheel Mouse, a trackball and a joystick were used in this study.

2.3 Procedure

Participants were randomly assigned the order of the devices used in the experiment to reduce the practice effect on cursor control performance. CAT automatically generated a set of specified pointing, clicking, double clicking and dragging operations based on the researcher's design. Each participant was asked to complete the whole set of the cursor control operations in a limited time. The accurate rate, movement time, cursor movement trajectory, and the speed of cursor movements were recorded in the database and were compared among three devices.

2.4 Data Analysis

Analysis of the recorded data involves comparing the accurate rate, movement time, cursor movement trajectory, and the speed of cursor movements among three different devices

3 Results

3.1 Clicking and Double Clicking

The participants were asked to perform left button single click, right button single click and left button double click on three different icon sizes: 0.5cm diameter, 1cm diameter, and 2 cm diameter by using three different pointing devices. The Joystick seems to be the most effective device for children with cerebral palsy to perform

clicking and double clicking. Table 2 indicated the percentage of successfully performing clicking and double clicking among three devices. The median value for the percentage of successful clicking by using the Joystick was 44.5 % compared with 5.5% of the Logitech mouse and 16.5 % of the Trackball. The similar results were shown on the median value for the percentage of successful double clicking by using Joystick was 100% compared with 16.7 % of Logitech mouse and 83.4% of the Trackball. Participants identified their difficulties in holding the mouse and press the right/ left buttons at the same time when using Logitech mouse. Joysticks and Trackballs are easier for children with cerebral palsy to perform clicking and double clicking since they do not need to hold the devices and press the buttons at the same time.

Table 2. The percentage of successfully performing clicking and double clicking among three devices

Participant	Logitech mouse		Trackball		Joystick	
	Click	Double Click	Click	Double Click	Click	Double Click
A	0%	0	0%	16.7%	56%	100%
B	56%	66.7%	78%	66.7%	56%	100%
C	0%	16.7%	0%	100%	89%	100%
D	0%	16.7%	11%	33.3%	0%	100%
E	11%	16.7%	22%	100%	11%	100%
F	67%	16.7%	22%	100%	33%	100%

3.2 Pointing and Selecting

Table 3 indicated the percentage of successfully complete the tasks of moving the cursor 10 cm and click in the various icon sizes: 0.5cm × 0.5 cm, 1cm × 1cm, and 2cm× 2cm. Table 4 summarized the speed when participants performing pointing and selecting by using three different devices. Based on the results of Tables 3 and 4, Joystick is most effective for participant A and participant C. Both of participant A and participant C did not successfully complete any tasks by using Logitech mouse and the Trackball. However, for participants B, D, and E, Trackball may be the first priority to use. Participant F may prefer the Logitech mouse to use. Each participant

Table 3. The percentage of successfully performing pointing and selecting tasks among three devices

Participant	Logitech mouse	Trackball	Joystick
A	0%	0%	56%
B	56%	78%	56%
C	0%	0%	89%
D	0%	11%	0%
E	11%	22%	11%
F	67%	22%	33%

Table 4. The speed performing pointing and selecting tasks among three devices

Participant	Logitech mouse (cm/sec)	Trackball (cm/sec)	Joystick (cm/sec)
A	N/A	N/A	0.54
B	0.87	1.25	0.87
C	N/A	N/A	0.61
D	N/A	0.42	N/A
E	0.99	1.62	0.85
F	1.65	0.90	0.63

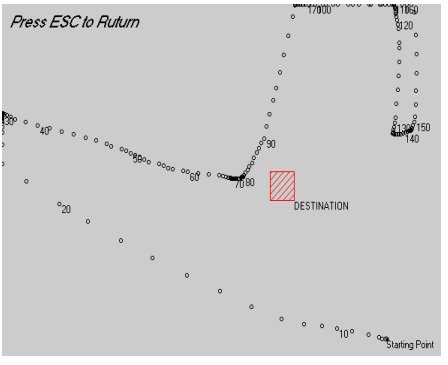


Fig. 2a. Movement trajectory of pointing and selecting for Participant A when using mouse

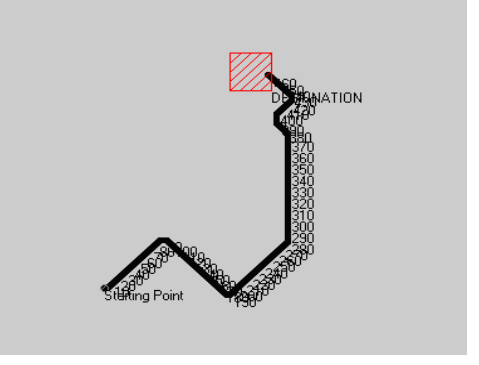


Fig. 2b. Movement trajectory of pointing and selecting for Participant A when using Joystick

performed differently in three devices. These results may due to their prior experience using cursor control devices. Figure 2a and 2b demonstrated the mouse movement trajectory of pointing and selecting when participant A operated mouse and Joystick. The path of mouse movement is much longer than that of joystick. In addition, participant A was not able to control mouse directly to the destination by overshooting the target area.

3.3 Dragging

Table 5 indicated the percentage of successfully complete dragging tasks by various cursor distance: 5cm and 10cm as well as in the various icon sizes: 1cm × 1cm, and 2cm× 2cm . Table 6 demonstrated the speed of performing dragging tasks among three devices. Joystick seems to be most effective for most participants. Participant A did not successfully complete any dragging task. Some research has reported dragging to be more difficult and error-prone than pointing. For participants B, C, E, and F, operating Joystick seem to be easier than other two devices when performing dragging tasks.

Table 5. The percentage of successfully performing dragging tasks among three devices

Participant	Standard mouse	Trackball	Joystick
A	0%	0%	0%
B	34. 38%	25%	46. 88%
C	0%	8.33%	62.5%
D	0%	3.16%	0%
E	31. 25%	6.25%	31. 25%
F	28.13%	15.63%	31. 25%

Table 6. The speed of performing dragging tasks among three device

Participant	Standard mouse (cm/sec)		Trackball (cm/sec)		Joystick (cm/sec)	
	5cm	*10cm*	*5cm*	*10cm*	*5cm*	*10cm*
A	N/A	N/A	N/A	N/A	N/A	N/A
B	0.53	0.60	0.75	2.21	1.47	2.47
C	N/A	N/A	N/A	N/A	0.50	0.73
D	N/A	N/A	N/A	0.54	N/A	N/A
E	0.35	0.49	N/A	0.69	N/A	0.64
F	0.65	0.79	0.68	1.95	0.38	0.73

The most common difficulty observed when performing dragging of persons with disabilities was in raising the mouse button in the correct position at the end of the drag. Figure 3a indicated the dropping error when participant B can start the dragging tasks by clicking the left button of the trackball but cannot release the button on the target area. Compared to Figure 3a, Figure 3b showed the movement trajectory of

Fig. 3a. Movement trajectory of Participant B performing dragging by using trackball

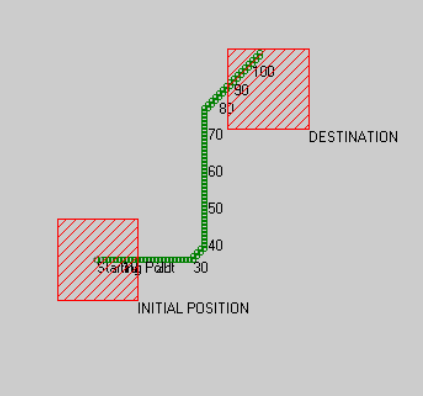


Fig. 3b. Movement trajectory of Participant B performing dragging by using Joystick

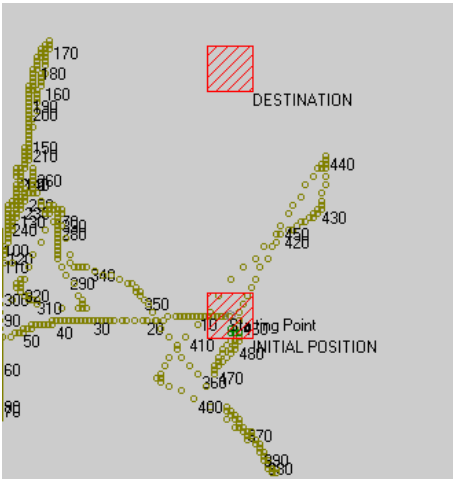


Fig. 4. Participant E cannot hold down the mouse button while moving the mouse

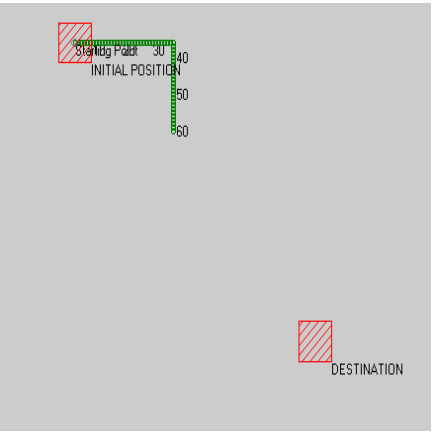


Fig. 5. Participant B get stuck when performing dragging tasks

dragging when participant B using joystick. The movement trajectory of joystick was smoother than that of trackball. In addition, participant B can successfully complete the dragging tasks.

Another common difficulty observed was persons with physical disabilities cannot not hold down the mouse button and move the button efficiently at the same time. Figure 4 shows the movement trajectory when participant E performing dragging tasks by using mouse. Sometimes, participants got stuck and had to abandon a drag which was shown in the figure 5.

4 Discussion and Conclusion

The results of this study indicated that Joystick seems to be the most effective for children with cerebral palsy to perform cursor control tasks. However, factors such as each participant's experience and preference may be related. The instrument use in this study, CAT, seems to be a convenient tool for rehabilitation therapists to assess clients' cursor control performance. Replication of this study using a larger sample size, or recruiting clients with different diagnosis and various ages should be developed.

References

1. Trewin, S., Pain, H.: Keyboard and Mouse Errors due to Motor Disabilities. Int. J. of Humana-Computer Studies 50(2), 109–114 (1999)
2. Hwang, F., Keates, S., Langdon, P., Clarkson, J.: A Submovement Analysis of Cursor Trajectories. Behavior & information Technology 24, 205–217 (2005)

3. Wu, T.F., Meng, L.F., Wang, H.P., Wu, W.T., Li, T.Y.: Computer access assessment for persons with physical disability: A guide to assistive technology interventions. In: Miesenberger, K., Klaus, J., Zagler, W. (eds.) ICCHP 2002. LNCS, vol. 2398, pp. 204–211. Springer, Heidelberg (2002)
4. Chen, M.C., Meng, L.F., Hsieh, C.F., Wu, T.F., Chu, C.N., Li, T.Y.: Computerized Assessment Tool for Mouse Operating Proficiency. In: Miesenberger, K., Klaus, J., Zagler, W., Burger, D. (eds.) ICCHP 2004. LNCS, vol. 3118, pp. 849–856. Springer, Heidelberg (2004)
5. Chen, M.C., Chu, C.N., Wu, T.F., Yeh, C.C.: Computerized Assessment Approach for Evaluation Computer Interaction Performance. In: Miesenberger, K., Klaus, J., Zagler, W., Karshmer, A.I. (eds.) ICCHP 2006. LNCS, vol. 4061, pp. 450–456. Springer, Heidelberg (2006)

User Modeling for Intelligent Interfaces in e-Learning

Ray Adams

CIRCUA, School of Computing Science, Middlesex University
London, United Kingdom
r.g.adams@mdx.ac.uk

Abstract. The emergence of the accessible knowledge society for all underlines the need for all to acquire the necessary knowledge and skills for inclusion. One way to do so is through e-learning, which itself should be accessible to all intended users. The mere provision of knowledge systems is not enough, since the need is for active and accessible learning that engages the participants effectively. This study explores the hypotheses that the solution is to be found in systems that: (i) support rather than replace the tutor, (ii) are accessible to the intended populations of users, (iii) can be adapted to the needs and individual characteristics of diverse users, (iv) are smart enough to adapt in real time to changing user needs, (v) reflect best practice in learning psychology, (vi) provide a high quality student experience, (vii) provide a high quality tutor experience and (viii) use valid student modeling. These hypotheses are evaluated through the five case studies. Accessibility and student modeling turnout to be the weakest points. All the other features are well represented in the case studies. None placed any kind of substantial emphasis upon accessibility. Only one of the case studies (case study 2; Cognitive Tutor Authoring Tools) makes a substantial effort in user modeling, being based upon the generic, cognitive model ACT. Even then, the focus tends to be on the typical or average user and does not address the problems of user diversity. Clearly, accessibility and user modeling need much more work in e-learning.

Keywords: user modeling, intelligent interfaces, e-learning, tutoring systems.

1 Introduction

The emergence of the accessible knowledge society for all underlines the imperative need for accessible ways to acquire the necessary knowledge and skills required for inclusion in it. The potential to push out such knowledge to all who need it requires the development of e-learning systems that are built upon the emerging accessible and smart information society technologies.

But, how can we create effective, stimulating e-learning applications? This study explores the hypotheses that the solution is to be found in systems that:

1. support rather than replace the tutor [20],
2. are properly accessible for the intended populations of users (teachers/tutors as well as students) [11], [12],

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 473–480, 2007.
© Springer-Verlag Berlin Heidelberg 2007

3. can be adapted to the needs and/or individual characteristics of different users [6],
4. are smart enough to adapt in real time to changing user needs [7],
5. reflect best practice in the psychology of learning [17],
6. provide a high quality student experience (interactive, rewarding and stimulating) [10],
7. provide a high quality tutor experience (content, delivery and results) [9] and
8. use a valid student modeling framework [2].

These hypotheses are evaluated through the consideration of five case studies of e-learning systems: KnowledgeNet [1], The Cognitive Tutor Authoring Tools (CTAT) [3], Open University of the Netherlands [8], SAGE (Self-Study Acceleration with Graphic Evidence) [5] and Advanced Joint English Teaching (AJET) [21].

2 Lessons from KnowledgeNet

KnowledgeNet is an e-learning developer of server-based courseware authoring and delivery system, supplying a range of Web-based applications [1]. Their applications include KnowledgeNet LIVE (a synchronous and real-time model) and KnowledgeNet INTERACTIVE (an asynchronous, interactive Web-based model for users who prefer self-paced learning methods).

KnowledgeNet support rather than replace the tutor. They have their own instructors and subject experts, focusing upon both human style interactions and valid contents selected by relevant human experts. The issue of accessibility for the intended populations of students is addressed by KnowledgeNet only indirectly. They talk about tailoring content to the users' specific learning needs and speed, though they do not yet specialize in accessibility of the interface or the requirements of special groups of users, those with disabilities or the older adult learners. However, it is clear that their learning materials can be adapted to meet the needs of different students. Whilst KnowledgeNet systems are not intelligent in the traditional sense of the term, they are designed to be smart enough to allow adaptation in real time, providing "write once, change anytime" facilities, being able to change contents "on the fly" with the underpinning use of a robust database engine.

Whilst, there is no overt commitment to the application of best practice in learning psychology [2], KnowledgeNet promote the importance of multimedia presentations and engaging the multiple senses of users, through sight, sound and simulation. This supports their commitment to provide a high quality student experience that is interactive, rewarding and stimulating. Technology, however powerful and cost-effective, is treated as necessary but not sufficient. They provide their own learning and content experts, to ensure a high quality tutor experience through high quality content and delivery. Though the systems can be adapted to meet the requirements of individuals or groups, the next stage of development is likely to be the use of valid student models that enable and support adaptive systems that maintain perceptual and cognitive accessibility to resources and content by adapting during run-time, reflecting user performance.

3 Lessons from Cognitive Tutor Authoring Tools (CTAT)

CTAT reflects a research and development program at the Human-Computer Interaction Institute, Carnegie Mellon University. Their focus is on the development of intelligent tutoring systems. Such systems can be very effective in enhancing student learning [14, 18], but are very hard to create and are used infrequently in universities and colleges, compared to simpler e-learning systems are much more popular. The CTAT authors estimate that 200 to 300 hours are required for one hour of instruction [4, 16, 19]. The aims of the CTAT program are to: (i) make tutor development more effective and (ii) provide research based insights into which features confer the best benefits and why. CTAT enables the generation of two types of tutor-systems: "Cognitive Tutors" and "Example-Tracing Tutors". Cognitive Tutors use a rule-based cognitive model and Example-Tracing Tutors that are a newer type of tutor-systems [14]. The latter provide the same tutoring functionality as Cognitive Tutors but have the added advantage that they do not require any programming [13]. The system requires examples of both correct and incorrect problem-solving behavior, to be recorded by the system-tool. The system generalizes and annotates these examples as a basis for Example-Tracing Tutors, or as a basis for guiding the development and testing of a cognitive model for use in a Cognitive Tutor.

Cognitive Tutors are based on the well-known ACT-R theory of cognition [4]. A cognitive model with production rules captures and compares actual student problem-solving behavior with the skills the student is expected to learn [13, 14]. The tutor-system tracks the student's performance and identifies problems. Producing a cognitive model for a Cognitive Tutor is a very difficult and time-consuming task that necessitates AI programming.

Example-Tracing Tutors are produced by demonstration rather than programming [13]. The author demonstrates to the system how students are required to solve each problem and the errors are expected. More problem authoring is required but no AI programming. CTAT provides three applications; first an external graphical user interface construction system based on the arrangement of interface widgets on a canvas by drag and drop but without programming, second a set of tools for task analysis, evaluating and modifying cognitive models, and third an editing system for cognitive models. Initial experiments find efficiency gains in system development time of 1.4 to 2.0. (Development time is halved.)

CTAT places equal emphasis upon the technology and human tutors. These system developers design their systems for a wide range of practitioners. System accessibility is not directly tackled by CTAT. They make use of cognitive models to address cognitive accessibility (access to content) but could consider sensory and perceptual problems. There is no focus on special groups of users, those with disabilities or the older adult learner but learning materials can be adapted to meet the needs of different students. CTAT systems are smart enough to respond in real time to changing student needs. Using ACT [4], the systems are closely related to relevant developments in cognitive and learning psychology and user modeling [2]. The emphasis is on creating worthwhile interactions for the students by creating smart systems. Their development of valid student models is an important contribution to this field.

4 Lessons from the Open University, Netherlands

Dewiyanti, Brand-Gruwel, Jochems and Broers (2004) [8] explored the student experience of asynchronous Computer-Supported Collaborative Learning Environments (CSCLE) at the Open University of the Netherlands, involving the responses of distance learning students, working in small groups. Student experience and satisfaction were recorded. Clearly the students valued the opportunity of working together and reported satisfactory experiences. The group project was an important contributor to group regulation and cohesion, requiring them to adopt active, constructive approaches to knowledge acquisition and achievement, more so than studying alone. Students felt that they had more control over their activities when working in a group.

This approach placed less emphasis on technology and more on collaborative working of distance-learning students, supported by asynchronous communication technology. This study supports collaborative learning over studying alone though there is no direct comparison. The systems used here were not particularly intelligent and the study did not include accessibility issues, despite the probability of students with disabilities selecting distance learning. The contents of their learning materials could be adapted to meet the needs of different students. They demonstrate the value of collaborative working and that is an important consideration for the use of e-learning.

5 Lessons from Self-study Acceleration with Graphic Evidence (SAGE)

Bell, Fonarow, Hays and Mangione (2000) [5] explored the SAGE system (Self-study Acceleration with Graphic Evidence) for web-based physician instruction on guidelines for care after myocardial infarction. The objective of their study was to construct and evaluate a Web application that would combine tutorial and statistical graphic simulation strategies in a self-study system. The need for this system is that many community physicians are unaware of important or new evidence that should their treatment of myocardial infarction (MI), even though printed leaflets are readily available. Such a system could be more effective as it can build on the student's existing knowledge and can interact with the user, demonstrating the known relationships between causes and effects. The former is a tutorial strategy and the latter is a simulation strategy. Both types of strategy were included in the design of this system. To test the efficacy of SAGE, twenty learning objectives for post-MI care were developed from nationally recognized guidelines. A pre-test and post-test was used based on one multiple-choice question per learning objective. (SAGE software was written with JavaScript, SQL, and Java.). The users entered their pre-test answers via a Web page, where they are directed to the guidelines related to the questions that they answered incorrectly. In addition, the system presented evidence from clinical trials by means of a Java applet providing using animated statistical graphics. Based on 158 participants of intended users, SAGE was compared with traditional learning methods (printed guidelines). Scores improved significantly from pretest to post-test

in both groups. Post-test scores were similar between groups, but the use of SAGE required less time studying, and greater satisfaction with learning efficiency.

The strengths of SAGE are to be found in the focus upon a well-defined group of intended users (community physicians working with post MI treatment in the community) by authors who were both subject experts and computer literate. They should be able to know their users well, even allowing for older physicians at some stage and able to consider themselves as having much in common with them. The issue of accessibility for the intended populations of students was not addressed explicitly, though personal awareness should not be ignored. It seems is clear that their learning materials can be adapted to meet the needs of the intended students. Whilst this system was not intelligent in the traditional sense of the term, it was designed intelligently in the face of user knowledge. This is a good example of a system that is designed for a niche Markey and one that will be refined further as additional experience is gained. As such, the niche product stands opposite the generic system.

6 Lessons from Advanced Joint English Teaching

Yang and Chen (2007) [21] provided a case study for what they call technology-enhanced language learning i.e. participant views of the use of web applications in language learning with senior high school Students (10th-grade, male students). They and their teacher were involved in technology-enhanced language learning (TELL) project in Taiwan known as Advanced Joint English Teaching (AJET). The students were given six Web-based tasks; group e-mailing, a Web-based course, an e-mail writing program, English homepage design, video-conferencing and chat room discussion. Whilst this study found that the AJET project demonstrated that the students liked and approved of learning English using the Internet, they had differing opinions about its benefits. But different learners had different perspectives. Passive users have need of more careful, pedagogical guidance than others. Self-directed learning turned out to be a critical factor in web-based foreign language learning.

Student reported different perceptions; some were involved more in the videoconferences than other AJET activities, whilst twenty percent did not see videoconferences as so helpful in improving language competency. Some were enthusiastic about chat rooms; others were not, citing problems with typing speed and poor writing proficiency. Individual students differed markedly in their e-mail communications and numbers of emails sent and received, thus altering their potential to learn. Some students reported that they were more creative when working with partners on-line. Satisfaction with the videoconference activity was lowest, followed by satisfaction with group e-mailing. Web page and presentation design had the highest satisfaction. They enjoyed variety of types of item available, text, graphics, animation, sound, voice, music, still pictures, and motion video, particularly video, audio and animation components of multimedia. The majority of students were most interested in personal web pages but film discussion, composition and email based pen pals generated lower interest scores even though they were seen by the same students has helpful for English language learning. This may be due, in part, to the high time demands made by the latter tasks. Conversely, for video-conferencing it was seen as interesting but least helpful for language skills development. They cited

problems with slow connection times, time lags, slow transmission speeds, screen freezing, and interrupted connections.

Most students reported strong improvements in vocabulary, reading, writing and listening skills, with smaller improvements in oral expression. They said that they had become more autonomous in their language learning and could independently find the meaning of difficult words or phrases. 85% of students reported that they had acquired new computer skills. 80% of students said that they had learned fields of knowledge they would not have otherwise been able to learn, like knowledge of the people, food, customs, culture, history, and resorts of other countries. Compared with traditional language learning styles of lecturing, 85% of students favored technology-enhanced language learning.

7 Overall Lessons

The present focus has been on the creation of effective, stimulating e-learning applications. This study explored the hypotheses that the solution is to be found in systems that:

1. support rather than replace the tutor,
2. are properly accessible for the intended populations,
3. can be adapted to the needs and individual characteristics of different users (tutors/teachers and students as well),
4. are smart enough to adapt in real time to changing user needs,
5. reflect best practice in the psychology of learning,
6. provide a high quality student experience,
7. provide a high quality tutor experience and
8. use a valid student modeling framework as well.

These hypotheses were evaluated through the consideration of five case studies of e-learning systems: KnowledgeNet [1], the Cognitive Tutor Authoring Tools (CTAT) [3], Open University of the Netherlands [8], SAGE (Self-Study Acceleration with Graphic Evidence) [5] and Advanced Joint English Teaching (AJET) [21] (see Table 1).

As discussed above and shown below, the five case studies are striking in the degree of diversity that they display, contrary to our initial expectations. At the onset, eight putative key features of an effective e-learning system were proposed but each of the case studies provided a different take on these features.

Table 1. Key features and case studies

Features	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
Case Study 1	√	-	√	-	√	√	√	-
Case Study 2	√	-	√	√	√	√	√	√
Case Study 3	√	-	-	-	√	√	√	-
Case Study 4	√	-	√	-	-	√	√	-
Case Study 5	√	-	√	-	√	√	√	-

The above table captures the diversity of the five case studies, from which it is clear that feature 2 (accessibility) and feature 8 (valid student modeling) are the weakest points (√ indicates a feature is covered well). All the other features are well represented in most of the case studies. None of these case studies place any kind of substantial emphasis upon accessibility considerations. Turning to user modeling, only one of the case studies (case study 2; Cognitive Tutor Authoring Tools) makes a substantial effort in that area, being based upon the cognitive model ACT [4] and aiming to be an intelligent tutoring system. Even then, the focus tends to be on the typical or average user and does not address the problems of user diversity. In conclusions, it is clear that, whilst most of the eight factors presented here are well treated in these five case studies, the achievement of accessible e-learning systems that make proper use of user models will need much more work.

References

1. A KnowledgeNet White Paper. Exploding the e-Learning Myth: Next-Generation, Web-Based Training Is Here Today and It Delivers the Wow Experience (2002) [Available Online] (viewed October 2006) http://www.knowledgenet.com/newsroom/whitepapers/elearningmyth.jsp
2. Adams, R.: Universal access through client-centred cognitive assessment and personality profiling. In: Stary, C., Stephanidis, C. (eds.) User-Centered Interaction Paradigms for Universal Access in the Information Society. LNCS, vol. 3196, pp. 3–15. Springer, Heidelberg (2004)
3. Aleven, V., McLaren, B.M., Sewall, J., Koedinger, K.R.: The Cognitive Tutor Authoring Tools (CTAT): Preliminary Evaluation of Efficiency Gains. In: Ikeda, M., Ashley, K.D., Chan, T.-W. (eds.) ITS 2006. LNCS, vol. 4053, pp. 61–70. Springer, Heidelberg (2006)
4. Anderson, J.R.: Rules of the Mind. Erlbaum, Hillsdale, NJ (1993)
5. Bell, D.S., Fonarow, G.C., Hays, R.D., Mangione, C.M.: Self-study from web-based and printed guideline materials. A randomized, controlled trial among resident physicians. Annals of Internal Medicine 132, 938–946 (2000)
6. Brusilovsky, P.: KnowledgeTree: A Distributed Architecture for Adaptive E-Learning. In: Proceedings of the Thirteenth International World Wide Web Conference (WWW 2004), pp. 104–113. ACM Press, New York (2004)
7. Christensen, J., Sussman, J., Levy, S., Bennett, W.E., Wolf, T.V., Kellogg, W.A.: Too Much Information. The Future of HCI. ACM Queue 4(6), 50–57 (2006)
8. Dewiyanti, S., Brand-Gruwela, S., Jochems, W., Broers, N.J.: Students' experiences with collaborative learning in asynchronous Computer-Supported Collaborative Learning environments. Computers in Human Behavior 23, 496–514 (2007)
9. Fletcher, J.: Evidence for Learning from Technology – Assisted Instruction. In: O'Neil, H., Perez, P. (eds.) Technology Applications in Education: A Learning View, pp. 79–99. Lawrence Erlbaum, Mahwah (2003)
10. Granić, A., Ćukušić, M.: Project UNITE - unified eLearning environment for European schools. Electronic Magazine for Promotion of IT in Education. Croatian Academic and Research Network, CARNet (2006)
11. Karim, S., Tjoa, A.M.: Towards the Use of Ontologies for Improving User Interaction for People with Special Needs. In: Miesenberger, K., Klaus, J., Zagler, W., Karshmer, A.I. (eds.) ICCHP 2006. LNCS, vol. 4061, pp. 77–84. Springer, Heidelberg (2006)

12. Kerscher, G.: The Essential Role of Libraries Serving Persons Who Are Blind and Print Disabled in the Information Age. In: Miesenberger, K., Klaus, J., Zagler, W., Karshmer, A.I. (eds.) ICCHP 2006. LNCS, vol. 4061, pp. 100–105. Springer, Heidelberg (2006)
13. Koedinger, K.R., Aleven, V., Heffernan, N.T.: Toward a rapid development environment for Cognitive Tutors. In: Proceedings of the 11th International Conference on Artificial Intelligence in Education, AI-ED 2003, pp. 455–457. IOS Press, Amsterdam (2003)
14. Koedinger, K.R., Anderson, J.R., Hadley, W.H., Mark, M.A.: Intelligent tutoring goes to school in the big city. International Journal of Artificial Intelligence in Education 8, 30–43 (1997)
15. Murray, T., Blessing, S., Ainsworth, S.E. (eds.): Tools for Advanced Technology. Learning Environments. Kluwer Academic Publishers, Boston, MA (2003)
16. Nash, S.S.: Learning Objects, Learning Object Repositories and Learning Theory: Preliminary Best Practices for Online Courses. Interdisciplinary Journal of Knowledge and Learning Objects 1, 217–228 (2005)
17. Schweizer, K.: An Overview of Research into the Cognitive Basis of Intelligence. Journal of Individual Differences 26(1), 43–51 (2005)
18. VanLehn, K., Lynch, C., Schulze, K., Shapiro, J.A., Shelby, R., Taylor, L., Treacy, D., Weinstein, A., Wintersgill, M.: The Andes physics tutoring system: five years of evaluations. International Journal of AI in Education, vol. 15(3), pp. 147–204. IOS Press, Amsterdam, Trento, Italy (2005)
19. Woolf, B., Cunningham, P.: Building a community memory for intelligent tutoring systems. AAAI 1987 pp. 82–89 (1987)
20. Woolf, B.: AI in Education. In: Shapiro, I., Stuart, C. (eds.) Artificial Intelligence – Encyclopedias, pp. 434–444. John Wiley & Sons, Inc, New York (1992)
21. Yang, S.C., Chen, Y.: Technology-enhanced language learning: A case study. Computers in Human Behavior 23, 860–879 (2007)

Ubiquitous Access to Learning Material in Engineering

Andreas Auinger[1] and Christian Stary[2]

[1] Upper Austria University of Applied Sciences, Department for e-Business, Steyr, Austria
[2] University of Linz, Department for Business Information Systems, Austria
andreas.auinger@fh-steyr.at, christian.stary@jku.at

Abstract. Advances in mobile computing allow new ways of any-time and any-where, networked and dislocated knowledge transfer. Such new ways of instant learning access are especially required in the field of engineering, due to its increasing complexity and diversification of tasks, and the resulting skill profiles of engineers. Qualification support has been achieved by providing an integrated e-/m-learning solution in the ELIE project (E-Learning In Engineering). It adapts didactically relevant content to both, the technical properties of multiple devices, and the situation of the learner. The structuring process for content is based on elicited experiences of coaches. In this way, it ensures high didactic value of learning material. Since this expert knowledge can be encoded into content directly, self-directed transfer processes can be triggered and managed effectively. The procedure for content development we propose enables coaches not only to follow the implementation of their knowledge in a transparent way, but also to reuse content, either as marked as didactic entities, redesigning them, or re-assigning didactic qualities to content. As technical enabler XML data structures have been developed for the encoding at the environment level. Both, in the web and the mobile learning setting, data are displayed sensitive to the user, the transfer situation, and the device at hand. In addition, users might link content elements to individually selected communication elements, such as entries in a discussion forum. The evaluation performed so far indicate several benefits for learners and coaches.

Keywords: mobile learning, didactics in e-learning, content management.

1 Introduction

By establishing the information society lifelong learning has become complementary to formal education, primarily driven by occupational needs. It is expected to qualify its members for a time of continuous changes, in particular with respect to skills they need to accomplish work tasks. Communication technologies and personal mobile devices enable access to information at any time at any location. Of particular interest in this context is the field of mobile learning (m-learning), as mobile networked applications might transfer knowledge effectively at the right time at the right location in the right way. Mobile learning is expected to facilitate lifelong learning [1], and leads to learning activities embedded in everyday (conversational) scenarios [2][3]. As mobile computing devices are available in different flavors, such as PDAs, smart

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 481–490, 2007.
© Springer-Verlag Berlin Heidelberg 2007

phones and portable computers, taking advantage of them has brought up a variety of facilities for interaction, navigation, and presentation of information. In addition, computer systems can be used in environments where they could not been used so far. Such a high degree of diffusion requires novel concepts, such as location awareness and context sensitivity. Taking into account user needs, work and situational context has to be considered as a prerequisite for m-learning, otherwise the challenges of achieving intelligent content delivery and engaging learning experience [4] cannot be met.

A recent survey study performed in the UK and Italy reflects the attitude towards m-learning of a total of 128 learners [5]. According to the study's results, mobile learning is considered unique due to its capability to enable learning effectively anywhere, anytime, *and* in a personalised way. Furthermore, mobile learning helps learners to identify areas where they need assistance and support. To that respect, adequately prepared content and highly self-contained, but motivating learning material could help, in addition to effective communication facilities. Tuned content and communication would allow focussed interaction within the peer group, and help contacting tutors and coaches.

Reflecting the findings so far m-learning can be considered a key advancement in the field of learner-centred knowledge acquisition and self-directed learnering. Recent approaches to m-learning stress the importance of a structured procedure to produce proper content, and suggest engineering-driven approaches for content production [6]. Unfortunately, these approaches lack concepts for identifying didactically relevant content elements in m-learning environments, as well as their context-sensitive specification and implementation. According to Dijkstra et al. [7], content structures need to meet didactic requirements to become Learning Objects (LOs). LOs are considered crucial for effective learning in distributed learning environments. Although current research in the field of advanced distributed learning stress the importance of modular content motivated by domain-specific didactics in academic environments [8, 9], LO development is driven by technical means and needs rather than didactic experiences and concepts. However, the didactic value of learning material is *the* decisive factor for online learning: "The material provided through online learning must have a high didactic value, since it is primarily intended for self-paced training" [10].

Developing material of high didactic value is a challenging task. As Kerres [11] and Euler [12] point out, bridging the gap between technology-centred development and didactically relevant content management, requires more than content structuring, namely the recognition of didactic expertise and situational parameters in the course of development: "Most of the efforts to support the preparation and deployment of accessible web-based learning material propose guidelines that prevalently address technical accessibility issues. However, little or no consideration is given to the didactic experts, and thus their didactic experience, in learning material development" [13]. Consequently, enabling multiple and dynamically generated representations of knowledge requires more than shifting content from conventional e-learning to m-learning environments: "Developing learning materials specially for mobile learning is better than re-using material developed to a PC".

In conformance to these findings our research has been targeting towards didactically motivated and effective content development for web-based and mobile

learning environments (cf. www.ce.scholion.at, www.mobilearn.at). In this paper we show in which way highly structured, didactically enriched content can be produced in a context-sensitive, but still efficient way. It has been effectively used in the field of electrical engineering, both in web-based e-learning and m-learning scenarios of use. As enabler a platform has been selected that allows for multi-channel content delivery on e-learning and m-learning devices. After detailing the didactics-driven content engineering process in section 2, we show the mapping of didactic elements to the XML content structures in section 3. In section 4 we briefly discuss the relevant features of the second generation e-learning platform SCHOLION that is required for the ELIE environment. Section 5 concludes the paper reflecting our results and indicating further research based on the practical experiences made so far.

2 Content Engineering Driven by Didactic Knowledge

The content engineering process for m-learning content comprises various steps. It is composed of 5 main steps: preparation phase, preliminary document analysis, structured interview, extended document analysis and mapping of didactic knowledge, and finally, the actual content authoring and delivery to a Learning Management System (LCMS). Different roles, including domain expert, didactic expert, author and LCMS administrator, should be involved in the process (cf. **Fig. 1**). The core process steps aim to identify how content can be handled conform to (domain-)didactic principles. The data to that respect are collected through interviews with domain experts. Subsequently, they are mapped to didactically enriched content elements. These three steps constitute the CoDEx method (Content made explicit from a didactic perspective) [14,15,16].

Preparation. In the course of preparation, the source materials for content development have to be identified. Based on their objectives and outlines, literature in the content domain and/or source content from existing courses are arranged to a concluding source document.

Eliciting Didactic Knowledge. The CoDEx method has been developed to make the relevant didactic knowledge of domain experts for content engineering explicit, and to map this knowledge to digital content. It requires the following steps.

Preliminary Document Analysis. Hereby didactic experts examine each source content. The levels of granularity, didactic components, such as definition, and means of orientation and navigation are identified. The level of granularity of the various sources may be quite different. Depending on the intended use of the content to be developed and the envisioned didactic scenarios, different levels of detail may be useful. Additionally, the didactic orientation and the didactic objectives of the source content should be examined. Finally, the following items should become transparent: (i) a reason to include the data into the digital content; (ii) relationships between the different sources of information; (iii) generic content types/elements/objects; (iv) alternative ways of navigation and navigational patterns.

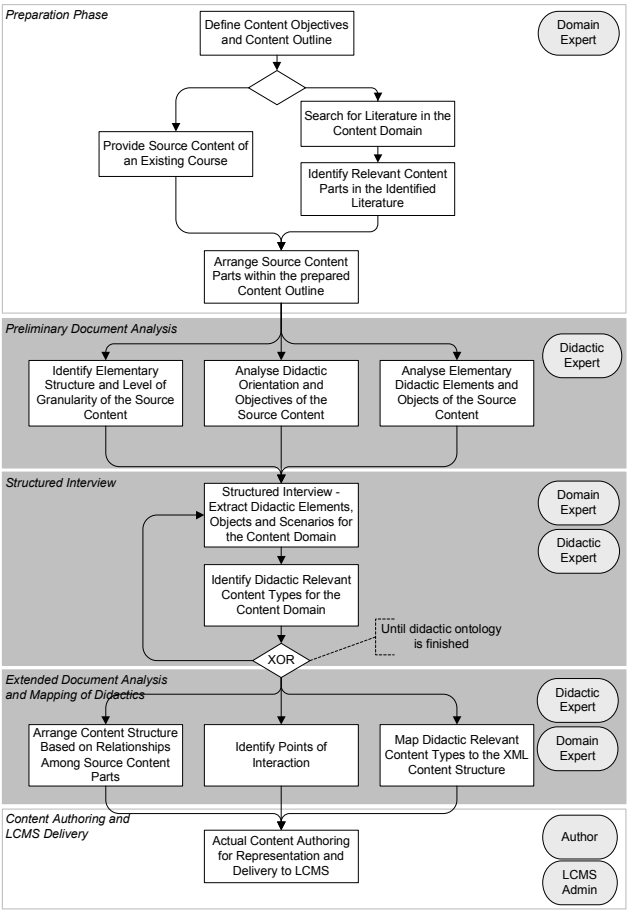


Fig. 1. The content engineering process (single iteration)

Structured Interview. According to Flechsig [17] each didactic entity needs to be well planned and requires a broad analysis of target group, learning culture, organisation, learning program, resources, demands, requirements, knowledge and competences. The most of these analysis steps also affect the content engineering process. In the course of a structured interview with the coach (or domain expert), the focal points of the didactic analysis will be treated from his/her perspective.

- *Organisation.* Organisational issues including content profiles, learner profiles and the organisational learning environment are clarified in this interview section.
- *Individual Positioning.* This section should make explicit individual approaches to knowledge transfer: Techniques to teach and individual didactic principles are explored.
- *Knowledge transfer.* This part of the interview deals with organisational activities during knowledge transfer activities, and the representation of relevant learning materials.

- *Communication*. The communication behaviour of the interview partner in the context of knowledge transfer is clarified. Conclusions about the social interaction and about social skills of the interview partner are drawn.
- *Technical support*. Current and future technical support of the knowledge transfer is clarified: Questions about the use of ICT-tools, the www or learning environments in conjunction with learning material are asked.

The structured interview clarifies individual, organisational and technical aspects of the knowledge transfer process. In the core part of the interview, didactically motivated elements such as didactic content types, interactive elements etc. are identified. Depending on the preparation and knowledge of the interview partner (domain expert, coach or teacher) these content elements and their interrelations can be balanced with the content elements in the learning materials.

3 Structure Mapping

In this phase, the elicited didactic elements and structures are mapped to an XML content structure.

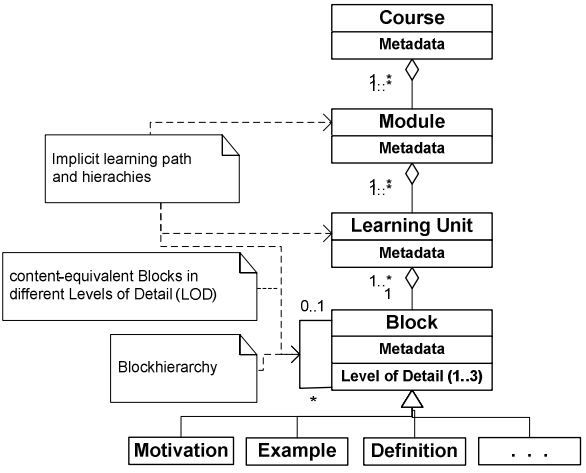


Fig. 2. Didactically motivated content structure

Content Structure. The content structure as described in Fig.2 comprises levels of details, domain-specific block types, and rendering levels for different device categories.

Levels of Detail. The LOD (Level Of Detail) concept is important for polymorph content delivery. It allows content developers to produce content on three different levels of granularity, and consequently, learners to retrieve differently detailed information of the same block type or on the same topic. A very common instantiation of that concept is to provide slides for presentation at LOD1, a text book structured according to block types at LOD2, and additional information or further material

(links, files in pdf format, videos and the like) at LOD3. At LOD1, learners might retrieve essential information at a glance with minimal effort in navigation. In combination with the filter function, mobile access can be designed as effectively as stationary access. In this way, access to content 'on-the-move' provides added value to knowledge transfer in virtual environments rather than burden users with cumbersome navigation activities.

Didactic Blocktypes. In the content area the content is grouped according to didactic principles, forming so-called blocks. Currently, about 15 generic block types have been defined and implemented. They comprise definitions, motivation, background, directives, examples, self test and other generic didactically relevant content structures. Additionally, some domain specific block types, such as source code, interaction etc. were defined for the field of engineering in the course of the ELIE project.

Rendering levels. Basically, content in the different Levels of Details can be used at various delivery platforms (PC, PDA and Smartphone). The content developer should only care about multimedia files and pictures. Hence, different display resolutions, different media players, and different capabilities of delivery platforms have to be considered and implemented in different, so-called rendering levels, which are also part of the content structure. Only specific combinations of LOD and delivery platforms have turned out to be effective for learning so far (cf. Table 1).

Table 1. Scenarios using content on different Levels of Detail and devices

Level of Detail	Granularity of Text	PC	PDA	Smartphone
LOD 1	presentation style	X	X	X
LOD 2	full text	X	X	
LOD 3	additional information	X		

Mapping of didactics to content structure. At this point, source content and didactic elements are mapped to the content structure jointly by the didactic expert and the domain expert. In the course of the ELIE project, the source contents were prepared in MS Word documents. The learning units, blocks, and block types were integrated in the document using comments and highlighted. Furthermore, interactive elements and adequate interactive discussion issues were inserted directly into the source content at identified points of interactions. Preparing source documents in this way was the precondition for the actual content authoring and delivery to the ELIE environment.

4 ELIE - Powered by SCHOLION

The ELIE environment is one of the instances of the SCHOLION platform [14], similar to mobiLearn (www.mobilearn.at). As a second generation educational environment it overcomes current limitations of distributed hypermedia environments for continuous user-centered teaching and active learning support. It provides individualisation facilities for knowledge producers and consumers, as well as a common information space for focused interaction.

The architecture reflects this open access concept based on a central data repository. The system tiers decouple the processing logic from data management and user interaction. The target user groups which are supposed to use the ELIE environment are knowledge consumers in companies, as well as trainers, lecturers and knowledge providers from educational departments, (business) schools and universities (knowledge producers). The benefits to be expected for both user groups are: Knowledge consumers are enabled to leave their passive role and become an integral part of the knowledge transfer process. This immersion of learners into the transfer process is achieved through telecommunication facilities and individual knowledge navigation and annotation.

Architecture. The technical architecture is based on 3 tiers, providing database-management-system independency for the data management, presentation and business logic (model, view and controller). Most instances of the platform use a Oracle 10g database capturing all data (including all media files and slides stored in BLOBs) and Tomcat 4.1.31. The application's servlet technology is based on Java 1.4 and generates a XML meta language, similar to HTML, but independent from styles and web-browser specific characteristics, that is rendered to XHTML 1.0, DOM 1.0 and CSS 1.0 (2.0). The architecture enables the personalisation of the graphical user interface as well as annotations and individual views.

Based on the model view controller paradigm, the architecture establishes a strict separation of presentation, business logic and data management by the use of XML and XSLT. The goal of MVC design is to separate the application object from the representation at the client-side and from the way in which the user controls it (Fig. 3). The architecture grants the following benefits:

- Re-use of model components: The separation of model and view allows multiple views using the same model.
- Enabling high design complexity
- Support for new types of clients
- Clarity of design
- Efficient modularity

Interaction facilities. The facilities of the ELIE environment have been decomposed into prominent areas that are relevant from a function and user perspective: The *learn space* has been designed as an easy-to-access content area to navigate within modules and to work within the learning units. It also comprises specific tools, such as a multimedia library, search features, and annotations. Using the annotation tool, users can individualise their content as well as mutually interact in a context-sensitive way - each content element might be directly linked to asynchronous or synchronous communication elements (chat, discussion forum entry etc.). Communication is supported by both synchronous and asynchronous tools. The chat forum and the Instant Messenger allow for synchronous communication within the platform, whereas the discussion forum and the information board support asynchronous communication. Both types of tools also enable users to collaborate for exchanging content or comments, and to form groups as they like. Finally, the office area is a set of personalisation features for each user. Users might create personal records, individual work spaces, and arrange their courses.

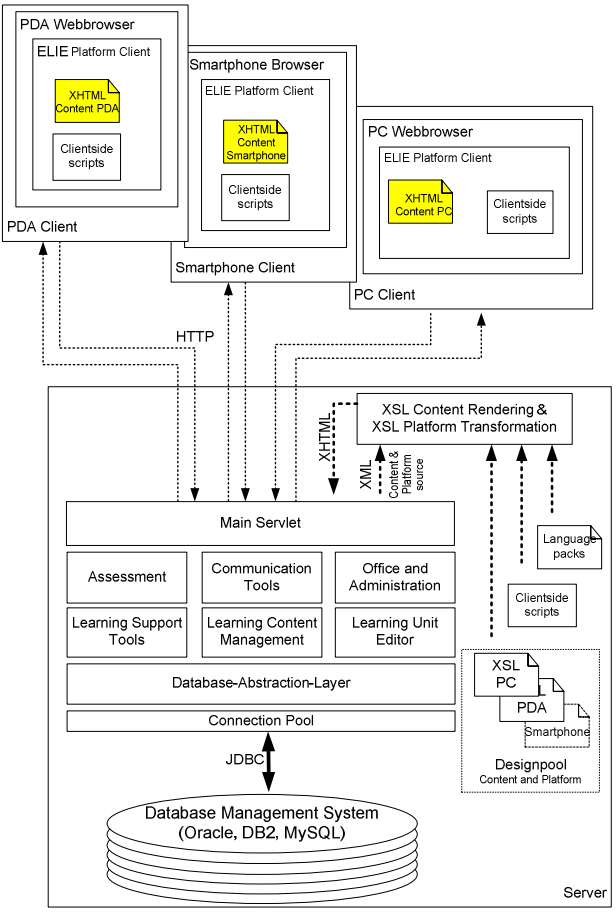


Fig. 3. ELIE Architecture for PC, PDA and Smartphone

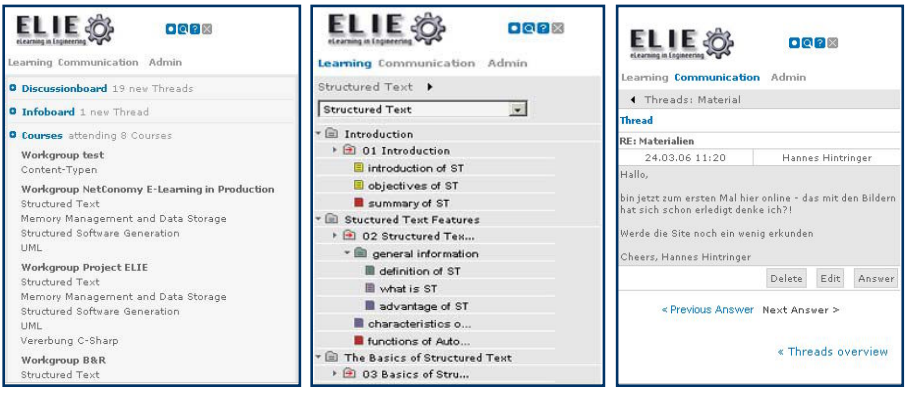


Fig. 4. ELIE PDA environment; left: startscreen; middle: content tree; right: discussions

The annotation tool enables the individual design of content and the learn space in the course of knowledge acquisition within a module. Learners can choose from a variety of functions: Individualization of textual content elements; Navigation within a module; Formatting and marking of text. Typical examples for individualisation are textual note taking, multimedia attachments, links to internal or external sources of information, underlining/ coloring of text, and direct links to entries in the discussion forum. In the platform annotations are stored in views that can not only be cascaded, but also transferred to other users (including coaches) or used as shared memories in work groups. The content figures display typical learning situations, both, at level of detail 1 and 2. All annotations are stored to user-individual views.

All main interaction spaces for knowledge transfer (content area with annotations, communications area and office area) are available for type of device (PC, PDA and Smartphone), taking into account the limited availability of space for displaying information and its manipulation, and the functionalities provided by various current vendors. The example in the figure shows content navigation using the PDA version. Further information on the PDA version can be found at [18].

5 Conclusions

Effective learning via mobile devices requires didactic preparation of content. Didactically typed elements and specific structures for multi-channel delivery on PC, PDA and Smartphone enable anywhere and anytime learning, information, navigation, and communication. The ELIE experiences in engineering have learners qualified in using novel forms of navigation and content-sensitive communication. Although high motivation and increased quality in knowledge transfer has been recognised so far, further empirical data have to be collected in a more systematic way.

References

1. Sharples, M.: The Design of Personal Mobile Technologies for Lifelong Learning. Computers & Education 34, 177–193 (2000)
2. Holzinger, A., Nischelwitzer, A., Meisenberger, M.: Lifelong-Learning Support by M-learning: Example Scenarios. ACM eLearn Magazine, vol. 5 (2005)
3. Sharples, M.: Disruptive Devices: Mobile Technologies for Conversational Learning (2002) http://www.eee.bham.ac.uk/handler/publications.asp
4. Beale, R., Lonsdale, P.: Mobile Context Aware Systems: The Intelligence to support Tasks and effectively utilize Resources. In: Proceedings of MobileHCI (2004)
5. Attewill, J.: Mobile Technologies and Learning. A little Update and M-learning Project Summary. Technology Enhanced Learning Research Centre (2005)
6. Vrasidas, C.H., Glass, G.V.: Distance Education and Distributed Learning. Information Age Publishing, Eurospan. London (2003)
7. Dijkstra, S., Jonassen, D., Sembill, D. (eds.): Multimedia Learning: Results and Perspectives. Lang. Frankfurt, Berlin, Bern, Brussels, New York, Oxford, Vienna (2001)
8. Schluep, S.: Modularization and Structured Markup for Web-based Learning Content in an Academic Environment. Shaker, Aachen (2005)

9. Taylor, J., Slay, J., Kurzel, F.: Learning Object Applications & Future Directions. Chapter in Learning Objects, Informing Science Press (2006)
10. Leidig, T.: L3 – Towards an Open Learning Environment. ACM Journal of Educational Resources in Computing, vol. 1(5) (2001)
11. Kerres, M., de Witt, C., Stratmann, J.: E-Learning. Didaktische Konzepte für erfolgreiches Lernen. In: Schwuchow, K.-H., Guttmann, J.: Jahrbuch Personalentwicklung & Weiterbildung 2003. Luchterhand (2003)
12. Euler, D.: Didaktik des computerunterstützten Lernens : Praktische Gestaltung und theoretische Grundlagen. In: Holz, H., Zimmer, G.(eds.) BW Bildung und Wissen. Verlag und Software. Nürnberg (1992)
13. Mirabella, Kimani, Catarci: A No-Frills Approach for Accessbile Web-Based Learning Material. In: Conference Proceedings of the WWW2004, New York (2004)
14. Auinger, A., Stary, C.: Didaktikgeleiteter Wissenstransfer. Interaktive Informationsräume für Lern-Gemeinschaften im Web. Deutscher Universitätsverlag, Wiesbaden (2005)
15. Auinger, A., Stary, C.: Effektive Content-Produktion für selbstgesteuerten, polymorphen Wissenstransfer. In: Ferstl, O.K., Sinz, E.J., Eckert, S., Isselhorst, T. (eds.) Wirtschaftsinformatik 2005 – eEconomy – eGovernment – eSociety, Physica-Verlag, Bamberg (2005)
16. Auinger, A., Auinger, F., Derndorfer, C., Hallewell, J., Stary, C.: Content Production for E-Learning in Engineering. In: Second International Conference on Interactive Mobile and Computer Aided Learning, Amman (to appear), (2007)
17. Flechsig, K.-H.: Kleines Handbuch didaktischer Modelle. Neuland, Eichenzell (1996)
18. Stary, C., Auinger, A.: Navigation Design for Mobile Access to Knowledge-Transfer Platforms. In: Salvendy, G (eds.) Human Computer International 2005. In: 11th HCII Conference, Las Vegas (2005)

ICT Educational Tools and Visually Impaired Students: Different Answers to Different Accessibility Needs

Stefania Bocconi[1], Silvia Dini[2], Lucia Ferlino[1],
Cristina Martinoli[2], and Michela Ott[1]

[1] Institute for Educational Technology- National Council of Research
Via De Marini, 6 16149 Genoa, Italy
[2] Istituto D. Chiossone for blind Onlus, Corso Armellini 11, 16122 Genova, Italy
ott@itd.cnr.it

Abstract. In order to guarantee equal opportunities to all students, the accessibility of ICT educational tools is worldwide considered a major issue. Nowadays, visually impaired students can take advantage of a large number of effective assistive technologies but, while using electronic material for learning purposes, they often encounter a number of different accessibility and usability problems. The variety of obstacles they may find on their way is quite large mainly because the term "visually impaired" encompasses a wide range of deficits, ranging from blindness to a number of other multifaceted, although less severe, visual impairments. The accessibility requirements for e-learning products established by the laws in force in the different European and non-European countries can be considered an important step onwards; further measures, nevertheless, are still needed to foster the actual "usability" of such products by sight impaired people.

Keywords: low vision, blind, accessibility, usability, educational software, e-learning.

1 Introduction

The Charter of Fundamental Rights of the European Union [2000] states that: "Any discrimination based on any ground such as sex, race, colour, ethnic or social origin, genetic features, language, religion or belief, political or any other opinion, membership of a national minority, property, birth, disability, age or sexual orientation shall be prohibited".

In the field of education, the basic concept of "Non – discrimination" entails the ability of all people to have "equal opportunity in education, regardless of their social class, ethnicity, background or physical disabilities" [Klironomos et al., 2005].

Students with disabilities have, then, the right to expect the same standard of education as their schoolmates and, in this view, they also have the right to access and use mainstream educational tools, including ICT based ones, which are generally referred to as "e-learning tools".

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 491–500, 2007.
© Springer-Verlag Berlin Heidelberg 2007

Such tools are worldwide considered powerful tools to foster learning [Hitchcock et al., 2003] but, at the same time, it is well known that, due to the widespread use of technological tools, "disadvantaged or excluded groups, including the unskilled, disabled and the elderly, face the danger of further marginalization […]", in fact, "with the advent of the digital computer, and its broad penetration […], disabled and elderly people face serious problems in accessing computing devices" [Stephanidis and Savidis, 2001]

Anderson [2006] underlines that, in the field of education, "while …technologies are beneficial and have been shown to help with educational tasks, their design and usability are an issue".

Students with disabilities may, in fact, face relevant difficulties both in "accessing and in "using" e-learning tools and, depending on the type of impairment, the types of obstacles encountered may vary considerably.

In the following, the issue of the accessibility and usability of e-learning tools by visually impaired students is tackled, taking into account the fact that in almost all the developed countries they represent a significant part of the school population.

In recent years, the World Health Organization [2004] has emphasized that "childhood blindness remains a significant problem, with an estimated 1.4 million blind children below age 15"; the problem appears even more relevant if we consider also the students with visual impairments other than blindness (the incidence of low vision being generally estimated three times greater than blindness). In the same report, in fact, it was found that, in 2002, there were 161 million (about 2.6% of the world population) visually impaired people in the world, of whom 124 million (about 2%) had low vision and 37 million (about 0.6%) were blind.

Data of the World Health Organization seem to confirm, then, that sight impaired students are a relevant percentage of the overall population of the students with disabilities [Viisola, M., 1999].

Such students, in principle, could highly benefit from using ICT for educational purposes but they actually, despite the availability of a growing number of technology-enhanced and sophisticated assistive devices, face a number of accessibility problems [Burzagli et al., 2004].

The task of listing all such problems is not easy, mainly because students with visual disabilities are a highly dishomogeneous category showing a wide range of different specific abilities, disabilities and needs; blind and low vision students, for instance, despite the fact that they are both often simply referred to as "visually impaired", present very different visual problems, find different obstacles and ask for different kinds of help and support.

What's more, even the category of low vision students is a highly heterogeneous one: it encompasses a wide variety of different visual characteristics. In this perspective, the different instructional tools (those ICT based, as well as the traditional ones), may meet or not to the needs of each single user, depending on her/his specific impairment.

After a brief review of the e-learning tools that are widely used for educational purposes, examples are provided of some problems encountered by students with visual disabilities in accessing and using e-learning material. Such problems often result into true obstacles for the effective use of the tools and may also have a negative influence on the overall learning process.

In the end, a glance is taken to the accessibility requirements addressing the specific needs of visually impaired people established by the Italian law in force, which directly recalls the Section 508 Rehabilitation Act[1].

2 Different E-Learning Tools for Different Visual Needs

While affording the choice of the e-learning tools to be used in concrete educational settings by visually impaired students, from one side, it is important to consider the nature, the specific features and the functionalities of the technological tools at hand; from the other side, it is necessary to take into account the actual, specific needs of the potential user/s (which are, of course, related to their impairments).

To this end, in the following, an overview and a basic classification of the main tools used in the field of e-learning is proposed; subsequently, a quick look is taken to the different needs of the different categories of visually impaired students and to the main obstacles that they may encounter.

2.1 What Do We Exactly Mean with the Term "E-Learning Tools"?

Anohina, [2005] defines "e-learning" as a learning process that "takes places via any electronic medium". In a global perspective, such a term refers, then, to any educational process making use of technological/electronic media and applications such as: "web-based teaching materials, hypermedia in general, multimedia CD-ROMs, web sites, discussion boards, collaborative software, e-mail, blogs, wikis, computer aided assessment, educational animation, simulations, games, learning management software, etc…" [Wikipedia]

In this view, the term "e-learning tools" encompasses at least those tools used for:

- ***"Online learning":*** those educational resources made available through inter-connected computer networks, comprising also synchronous and asynchronous communication tools, when used in an educational perspective.
- ***"Computer-based learning":*** those learning materials locally available on the user's PC and used when the computer is not connected to a network.
- ***"M-learning":*** those educational tools made available through "mobile devices" such as palmtops (or handhelds), Personal Digital Assistants (PDAs), tablet PCs, mobile or smart phones; such tools, may also take advantage of the connection to the net via "wireless transmission" [Hoppe et al, 2003].

The concept of "e-learning tool" is, then, linked both to the media (hardware devices) employed and to the programs (software applications) used to support the educational process. Such software applications can be roughly divided into:

- ***E-learning platforms:*** those internet-based environments expressly addressed to the delivery of integrated electronic educational contents and to the management of a variety of educational activities aimed at fulfilling specific educational objectives

[1] Section 508 of the Rehabilitation ACT Subpart C - Functional Performance Criteria § 1194.31).

[Lin and Kuo, 2005]. All the digital contents made available by and through such platforms, are generally called "learning objects".

- *Web based applications:* those applications (both designed for educational purposes and used to fulfil educational objectives) which are directly accessible using any available browser and which don't need to be installed on the user PC.
- *Standalone applications:* those products (both "educational" and "used for education") which cannot be used directly via browser but that need to be installed locally, on the user machine; this category includes also products "downloadable" from the Web, but that still need to be installed on the computer.

2.2 E-Learning Tools and the Needs of Visually Impaired Students

Examples of what kind of accessibility and usability problems can be found in the different categories of software applications are provided in the following, keeping aside, for the moment, all the possible problems linked to the use of computers in their standard configuration and of other specific hardware devices.

While considering such obstacles, it is important to reflect on the fact that they are strictly related to the type of user impairment. Blind and low vision students encounter different types of obstacles: in order to fully to access the contents, in fact, the first category needs necessarily to rely on screen readers, while the second category, thanks to optical aids and/or to specific customization options, may access a much wider variety of software applications, including, often, those with graphic interface. Other significant differences can be found among the needs of the different categories of low vision students due to the wide variety of their visual impairments [Dini et al, 2006].

E-learning Platforms and the Needs of Visually Impaired Students: Some Examples. E-learning platforms often pose a number of different problems to visually impaired students: small characters, crowded pages, pop up windows, iconic menus, complex forms to be filled in, etc…The availability of appropriate customizing options could help to avoid such problems.

Fig. 1 shows the main menu of an educational platform[2] used in the field of hospital education; on the left side of the following figure one can see the standard configuration, while, on the right side, one of the possible "personalized" configurations is presented. This platform cannot be used by blind users because it is fully based on visual presentation and no screen reading facility is available.

Such a platform, on the contrary, presents a number of facilities allowing personalization and this may help other categories of visually impaired students; the right side of fig. 1 shows, in fact, that the magnifier may significantly help vision and it also proves that menu objects can be moved around the screen thus meeting, for instance, the specific needs of students with reduced visual field or with tunnel vision.

Such customizations are linked to the user's profile and therefore they are maintained for any subsequent access to the platform made by the same user.

Web Based Tools And The Needs Of Visually Impaired Students: Some Examples. Web based tools generally do not cause excessive problems to blind and low

[2] http://scuolainospedale.indire.it

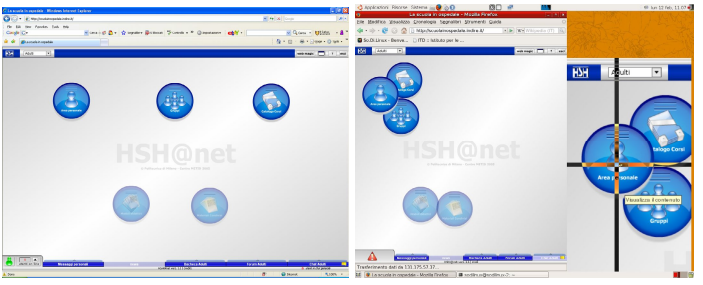


Fig. 1. Menu of an e-learning platform in standard (left) and personalized configuration (right)

vision users since they generally follow the recommendations of WAI[3] and W3C. Fig 2 shows an example of educational web based application (namely one of the tests used for obtaining the European Computer Driving Licence[4]) which can be profitably used by blind and also by a wide range of low vision students.

Fig. 2. Online educational test in standard (left) and personalized configuration (right)

Such an online-test can be fully keyboard driven, it is compatible with screen readers and supports magnifiers (right side of the image); it thus meets the needs of those users that need screen reading facilities as well as the needs of those who require screen enlargement; other categories of low vision students, nevertheless, (namely those with reduced visual field) may still have problems in using it.

Stand Alone Applications and the Needs of Visually Impaired Students: Some Examples. Educational software products are a very large category showing a wide number of different technical solutions; in the last years, GUI (graphic user interfaces) have been widely used and also the mouse has been often used as the only input system; recently international and national initiatives have produced significant changes in the direction of producing more accessible products even in this field.

[3] Web Accessibility Initiative (WAI) http://www.w3.org/WAI/; The International Webmasters Association (IWA) http://www.iwanet.org/

[4] ECDL example tests: http://www.webquiz.it/showquiz.asp?ABCDETRL accessed February 2007.

Fig. 3 shows a screenshot of ALFABET[5] an educational stand alone product devoted to mastering abilities in dictionary search and alphabetical ordering. The task at hand is to complete the letter sequence following the correct alphabetical order, picking up the missing letter among those available at the bottom of the screen. This task which, in principle, can be useful for both blind and low vision users can actually be performed only by partially sighted students. In fact, even if the blinds can be supported by a screen reader (and the related pointer emulating mouse functions), they cannot fully access the program contents: they will, in fact, miss the overall position of the objects in the screen and this will result in a real obstacle for the task comprehension, for the possibility of interaction and, finally, for the achievement of the underpinning educational objectives.

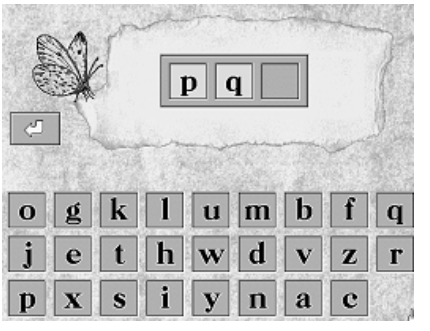


Fig. 3. Educational standalone software application: example of an exercise which does not meet the needs of blind students

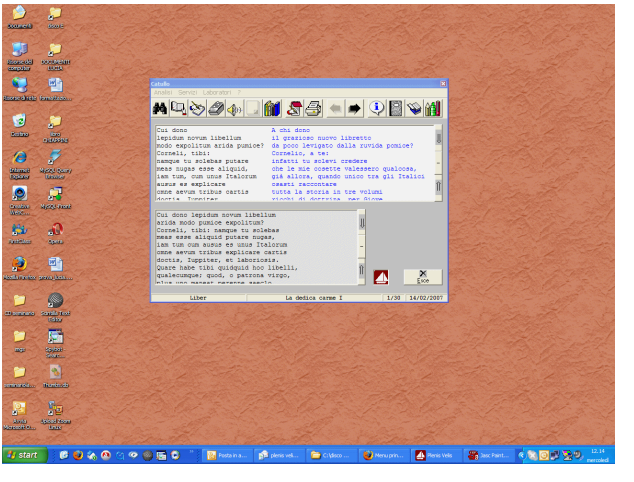


Fig. 4. Educational standalone software applications: example of a product which meets the needs of blind students but cannot be fully used by low vision students

[5] ALFABETH – (1997), Dainamic software accessed February 2007 http://www.dainamic.be/

Fig. 4 shows a screenshot of Plenis Velis[6], an educational standalone product devoted to the learning of the Latin language.

Such a commercial product perfectly supports the basic screen reading facilities thus meeting the needs of those students who need to rely on speech synthesis; such a product, on the contrary, can hardly be used by most low vision students mainly because it is completely text based, characters cannot be enlarged, the screen is very crowded, the screen partition (which is meaningful and supports task comprehension) is confusing and results in a very scarce legibility of the whole page.

In addition, full screen facilities are not allowed and the program window, which occupies a very small part of the desktop screen, cannot be enlarged (as shown in figure 4); to this end, changing the screen resolution doesn't help and, in such conditions, the eventual use of screen magnifiers doesn't help very much as well.

3 The Needs of Visually Impaired Students and the Laws in Force

The accessibility requirements for e-learning products established by the laws in force in the different European and non-European countries can be considered an important step onwards, in that software producers should behave consequently and we reasonably can expect that, in the future, they will deliver products better meeting the needs of users with disabilities.

Coming back to the point of view of visually impaired subjects, can we say that their needs are fully addressed by existing regulations and laws? To what extent? Are there any other "usability" aspects that should be considered?

In the following, while trying to answer these questions, we refer to the Italian Law in force 4/2004[7], which is directly founded on Section 508 of the Rehabilitation Act of the US Federal Government.

The correlations between the law requirements and the type of disabilities are not made explicit neither in Section 508 nor in the Italian law; even so, it is clear that the majority of the actual provisions address the needs of visually impaired people, as shown in Table 1.

Summarizing we can say that a great majority of the needs of visually impaired students is covered by the law provisions: alternative keyboard navigation is required in that it is essential for people with visual impairments who cannot fully rely on pointing devices, such as a mouse (Req. 1); the preservation of the basic accessibility features provided by operating systems is requested, including the maintenance of a number of features which are fundamental for low vision users such as the possibility of enlarging characters, reversing the color scheme, setting color brightness and contrast, increasing the contrast between text and background (Req. 2); other possible display selections made by the user (e.g. those related to the colors and the general layout of the screen) should be maintained as well (Req. 8); complete support for screen reading tools is required both in the case of graphic elements and of written texts in that this helps not only to allow the full comprehension of screen contents but also to make possible the written interaction (test answers, written commands, choices etc…)

[6] Plenis velis – Plenis Velis, a cura di Lucio De Rosa, Cappelli Editore 2001.

[7] LAW 4/2004 "Provisions to support the access to information technologies for the disabled" accessed February 2007, http://www.pubbliaccesso.it/normative/law_20040109_n4.htm

Table 1. Italian law 4/2004: main requirements for educational software applications and their basic correlations with the two main categories of visual impairments

Italian Law 4/2004 requirements[8] (with reference to the corresponding article of Section 508)	BLIND	LOW VISION
Req. 1. The functions provided by the user interface must be able to be activated by means of keyboard commands in cases where a description of the function or the result of executing it can be provided. Section 508: 1194.21 (a)	•	•
Req. 2. Commands and functionalities of the user interface must not limit or disable the accessibility functionalities of the operating system, made available by the manufacturer of the operating system. Section 508: 1194.21 (b)	•	•
Req. 3. The application must provide sufficient information, such as identification information, operations possible and status, on objects contained in the user interface so that the assistive technology can identify them and interpret their functionalities. Section 508: 1194.21 (d)	•	•
Req. 4. Where graphic symbols are used to identify controls, status indicators or other programming elements, the meaning assigned to these symbols must be consistent within the framework of the whole application, including the user interface. Section 508: 1194.21 (e)		
Req 5. Textual information must be provided using the functionalities of the operating system provided to display text. Textual content, the location of the insertion point and the text attributed must be available. Section 508: 1194.21 (f)	•	•
Req 6. Applications that use audio signals must provide an equivalent visual functionality, in accordance with any conventions of the operating system. Section 508: 1194.31 (c)		
Req 7. Animations, graphic or sound elements and differences in colour must not be used alone to provide information or indicate or request actions. Section 508: 1194.21 (i) (h)	•	•
Req 8. Applications must not overlap selections made by the user with regard to contrast or colour levels and other display attributes. Section 508: 1194.21 (g)		•
Req 9. The user interface must not contain flickering text elements, objects or other elements with a intermittent frequency greater than 2 Hz and lower than 55 Hz. Section 508: 1194.21 (k)		•
Req 10. The active "focus" element of a user interface must be clearly identifiable. Identification and variation of the focus must be indicated at the application programming interface (API) level, so that the assistive technology can manage them. Other elements that require action by the user must also be adequately indicated. Section 508: 1194.21 (c)	•	•
Req 11. Supporting documentation for the product and accessibility characteristics must also be made available in accessible electronic form. Section 508: 1194.41	•	•

between the software application and a user who, for instance may require a speech synthesizer (Req. 3 e 5); picture, photos, animations should always have non- graphic equivalents: they should not be the only way to convey information, since they represent a concrete problem for the visually impaired (Req. 7); elements which may

[8] English translation of the Italian law requirements following DM080705, accessed February 2007 http://www.pubbliaccesso.it/normative/DM080705-D-en.htm

disturb vision and/or reading, such as high – frequency flickering elements should be avoided (Req. 9); the location of the focus point should be easily detectable both by human eye and by assistive technology (Req. 10); documentation and other material supporting the use of software applications should be available in digital format, so that they can be fully read by any kind of users, including those with visual impairments (Req. 11).

4 Conclusions

The paper has tackled the issue of the accessibility/usability of educational material by visually impaired students. While examining a few educational products, the fact that different types of visual impairments have different needs and ask for different solutions has been underlined. A quick overview of the main provisions established by the Italian law in force has been taken and, as a result, it has been shown that they address a great majority of the needs of visually impaired users, establishing a significant number of "accessibility" features to be met.

The important issue, of their "usability" has been poorly addressed, instead, and we are aware that it needs further specific reflections. To this end, in fact, not only the interface features but also the product's basic functionalities should be considered and examined in a cognitive perspective. The idea of "usability" is, in fact, directly linked to the degree to which it "takes into account the human psychology and physiology of the users, and makes the process of using the system effective, efficient and satisfying"[9].

These functionalities are, of course, linked to the product's user interface but mainly depend on the contents as well as on the structure and the educational strategy at hand.

References

1. Charter of Fundamental Rights of the European Union Article 21-1 Non-discrimination (accessed February 2007) http://ec.europa.eu/comm/external_relations/human_rights/doc/charter_364_01en.pdf
2. Klironomos, I., Antona, M., Basdekis, I., Stephanidis, C.: EDeAN Secretariat for 2005 White Paper: promoting Design for All and e-Accessibility in Europe Universal Access in the Information Society International Journal vol. 5(1) (June 2006)
3. Hitchcock, C., Stahl, S.: Assistive Technology, Universal Design, Universal Design for Learning: Improved Learning Opportunities Journal of Special Education Technology, vol. 18(4) (2003)
4. Stephanidis, C., Savidis, A.: Universal Access in the Information Society: Methods, Tools, and Interaction Technologies, Universal Access in the Information Society, vol. 1(1). Springer, Heidelberg (2001)
5. Anderson, P.: The Future of Human-Computer Interaction in Emerging Technologies for Learning, BECTA (accessed February 2007) (2006) http://becta.org.uk/corporate/publications/documents/Emerging_Technologies.pdf

[9] Wikipedia "User Interface" http://en.wikipedia.org/wiki/User_interface accessed February 2007.

6. World Health Organization, WHO, Magnitude and causes of visual impairment Fact Sheet N°282 (accessed February 2007) (2004) http://www.who.int/mediacentre/factsheets/fs282/en/print.html
7. Viisola, M.: Statistics on Children with Visual Impairments ERIC Education Resource Information Center Publication ED432106 (accessed February 2007) (1999) http://eric.ed.gov/
8. Burzagli, L., Emiliani, P.L., Graziani, P.: Accessibility in the Field of Education, Lecture Notes In Computer Science. In: Stary, C., Stephanidis, C. (eds.) User-Centered Interaction Paradigms for Universal Access in the Information Society. LNCS, vol. 3196, pp. 235–241. Springer, Heidelberg (2004)
9. Anohina, A.: Analysis of the terminology used in the field of virtual learning Educational Technology & Society, vol. 8 (3), pp. 91–102 (2005)
10. Wikipedia, The free Encyclopedia - Electronic learning (accessed February 2007) http://en.wikipedia.org/wiki/Online_Learning
11. Hoppe, H.U., Joiner, R., Millard, M., Sharples, M.: Guest editorial: wireless and mobile technologies in education". Journal of Computer Assisted Learning 19, 255–259 (2003)
12. Lin, C., Kuo, M.: Adaptive Networked Learning Environments Using Learning Objects, Learner Profiles and Inhabited Virtual Learning. In: Worlds Proceedings of the Fifth IEEE International Conference on Advanced Learning Technologies (ICALT'05) (2005)
13. Dini, S., Ferlino, L., Gettani, A., Martinoli, C., Ott, M.: Educational software and low vision students: evaluating accessibility factors. Universal Access in the Information Society, Springer Berlin / Heidelberg, ISSN 1615-5289 (Print) 1615-5297 (Online) (2006)

Modeling Content and Expression of Learning Objects in Multimodal Learning Management Systems

Giorgio Brajnik

Dip. di Matematica e Informatica,
University of Udine, Italy
giorgio@dimi.uniud.it

Abstract. The paper discusses the relation between accessibility and multimodality of learning objects. I present a framework, rooted in linguistics, that supports a clear distinction between content and expression of learning objects. The framework supports different types of content elements, and different expression elements in which content can be rendered, which include different expression media, multimedia components and spatial/temporal/hyper-relationships. I show what accessibility means in such a framework, and what the consequences of accessibility are in terms of automatic transformations of expression that can occur on student platforms. I argue that accessibility of learning objects, though making them multimodal, is not a satisfactory solution from the viewpoint of learning effectiveness. I claim that content modeling needs to be a crucial step in the development of any learning object, and propose notations based on conceptual maps and UML diagrams.

Keywords: knowledge modeling, accessibility, operability of user interfaces.

1 Introduction

One function of Learning Management Systems (LMS) is to deliver one or more learning objects (LO) to students that access the LMS from various user platforms and in different situations. A student might access the LMS at home through a PC with a graphical browser; another from her PDA while sitting in a train; a third one using a PC at the university library, with audio turned off; a blind student might use a screen reader in a classroom during a lecture; a fifth one might have selected a user profile so that podcasting is her preferred way to access the instructional material; and a sixth one usually prefers to print lessons on a low-resolution and slow ink-jet color printer.

Other functions of LMS are to permit students to solve exercises and submit projects, to access collaborative features of the LMS (like a chatroom, a forum, a blog to view others' postings or to contribute with new content, to get in touch with other students, tutors, instructors), to access help/support sections of the LMS, or finally to use administrative services (to enroll to a course, to withdraw from a joint project, to view one's own ratings, etc.). In all these cases, students may access the LMS with a variety of devices, in a variety of situations.

In all these cases designers (of the LMS and especially authors of LOs) have no control on the platform adopted by the student to use the learning material. Yet, in all

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 501–510, 2007.
© Springer-Verlag Berlin Heidelberg 2007

these cases students should experience high levels of quality, especially in terms of usability, accessibility, engagement, memorability, and learnability [7]. In fact, learning effectiveness of LOs is heavily affected by motivational and persuasive factors embedded within LOs. Also the LMS may affect motivation, persuasiveness and learning effectiveness: usability mishaps, accessibility barriers, media interference within LOs and/or LMS may dramatically reduce teaching effectiveness. Furthermore, not every media is equivalent to the other ones with respect to learning effectiveness: one could conceive using podcasting for an introductory lecture on operating systems, but wouldn't use it for a lecture on Cauchy's calculus theorem.

Very often multiple media are used to enhance user experience, sometimes in a redundant way (like adding sound effects associated to user inputs), sometimes to complement each other (like using a visual simulation of electrical current in addition to a textual description of the phenomenon being simulated), sometimes as alternatives (like adding textual captions to a video for the benefit of those who cannot see the images).

What is the extent to which LOs can be adapted (or made adaptable) to these different learning situations? What kind of guidelines need to be given to LOs authors in order for them to provide redundant information? Is accessibility enough?

In this paper I explore these questions, laying a foundation upon which to look to issues like accessibility, multimedia, multimodality of user interfaces and LOs. An important conclusion is that content, defined in a sounder way than what is current understanding for websites and applications, has to be modeled separately from expression. I propose practical ways to isolate and model it.

2 Content Versus Expression

It is important to be able to separate content from expression within a LO (actually, the same applies also to the LMS hosting the LOs) in order to be able to clearly distinguish two important types of questions: those dealing with the meaning that has to be conveyed by the LO, and those dealing with how meaning is to be expressed. If a clear separation can be achieved, then it would simplify decisions regarding what to say, which arguments to draw upon in order to be convincing, which level of detail should the LO reach, what kind of external references to include, etc. It would simplify also decisions on the different ways in which the *same* meaning can be expressed, which media and combination of media would be more effective, and how the LO would be used (using different media and different modalities).

We base our distinction between content and expression on a linguistic approach, and assume that a LO can be thought of as a communication artifact (a message), involved in communication processes that occur between the LOs' authors and students. Additional details are available on [2]. Hjelmslev [6] suggested to consider a message as the combination of content, which corresponds to the meaning being conveyed by the message, and expression, which corresponds to the ways in which the meaning is expressed so that it can be perceived by human beings. Both content and expression can be discussed in terms of substance and form.

2.1 Content of a Learning Object

The content of a message is the result of a process of sensation-perception-interpretation guided by the context of the addressee (in the case of a LO the addressee is the student). By *context* I mean a combination of the physical situation the user/student is in (physical environment, input/output interaction devices, posture, light and noise conditions), his/her goals (final/behavioral/life-time goals [5]), beliefs, attitude and motivations.

Substance of content, for a generic LO, can be classified in the following types.

- *Information of interest*, which corresponds to the primary information that is being conveyed by the LO, and is what a student would expect to see/read/hear/learn to satisfy his/her final goals. Concepts belonging to the domain of interest, explanations, questions, answers, clarifications, examples, counter-examples, descriptions of phenomena are all information of interest. Also help material, tooltips and usage instructions belong to this category, as well as (external or internal) promotional messages.
- *Bearing information*, which supports the user/student in navigating and interacting with the system. Bearings can be further classified as:
    - *Location information*, that provides clues regarding the current position in the virtual space implemented by the LOs and the LMS. Examples are breadcrumbs [7], sentences like "question 3 our of 6", the lecture or page title, the title of an inset in the page, the label "on" attached to an animated switch which is part of a simulated electric circuit in the LO.
    - *Direction information,* which describes possible destinations that can be reached, steps that can be performed or activities that can be started. Examples include labels of links/buttons for moving among pages, menu entries, navigation bar entries, labels of buttons to control a simulation.
- *Access information*, used to activate functionalities of the system and achieve some response. To enable choice, very often access information is coupled to direction information. Access information can be split in:
    - *Sequential access schema*, that support a linear scanning of content units that are sorted according to some pre-defined criterion. Typical examples are the previous/next pairs of buttons that support scanning of pages of a tutorial.
    - *Indexed access schema*, that support access to units in any order. Examples include navigation bars, list of links, single buttons/links, drop-down menus.
    - *Filtered access schema*, that support operations for restricting the set of possible destinations. The most typical example is in a faceted classification interface [7]; other examples are forms used to select lectures, exercises, or postings in a forum.
- *Functional information*, that has to be provided by students in order to complete certain tasks. This includes, for example, answers to exercises.

Form of content refers to possible relationships that can be defined among content units. These relationships usually derive from a taxonomy or ontology of the domain: when certain classification criteria are used (f.e. by subject, by audience, by function,

etc.)[10], appropriate categories are identified and proper relationships among them are defined (f.e. “type-of”, “instance-of”, “cause-of”, etc.). Other relationships can be identified by adopting the Rhetorical Structure Theory (initially applied only to text [9], but more recently also to multimedia objects [3]). This theory can be used to determine if content units are consistent by identifying patterns of relationships that should hold among content units (f.e. unit A *contrasts* unit B, A *is an example of* B, A *provides details for* B, A *explains* B, etc.). Finally, content form includes also abstract task models, specifying combinations of steps (like sequences, loops, conditionals, fork/join points) that need to be followed in order to achieve the task goal.

Consider the example shown in Figure 1. From the content perspective (i.e. by not considering expressive aspects), its substance includes several location information providing an overview of the page (f.e. “The Magnetic Circuit”, “Magnetic Force”), location information (f.e. the knob position set to 5v, the display on the switch unit displaying “5V”, etc.). Except for the knobs that can be turned, this page does not include any direction information (buttons on the bottom of the page, labeled with “back” and “next”, do not describe any destination; they would if they were labeled with the name of the destination, like “magnetic flux” rather than “next”). The content includes also concepts like magnetic flux, voltage, currents, magnetomotive force, switch units, voltage and current measuring units, current measuring units, lamps, electrical circuits, series connection, etc.; these constitute the information of interest.

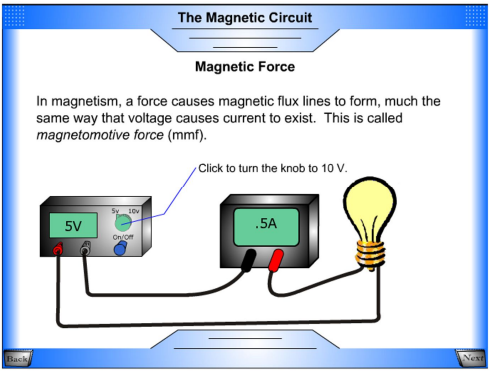


Fig. 1. A screenshot of a scene of the LO on magnetic circuits developed by T. Bartelt and J. Janikowsi, 2006 and published by Wisc-Online, www.wisc-online.com/objects/index_tj.asp ?objID=DCE14906

The form of the content of this example includes several relationships. Some are expressed in the text introducing the concept of magnetomotive force: we can see that authors adopted the magnetism-as-electricity metaphor, which is a relationship. Other relationships concern the electromagnetic phenomena being shown by the simulated circuit (f.e. when the button is set to “on” the lamp turns on; when the voltage knob is set to 10v, the lamp gets brighter). While the top part of the screenshot introduces the metaphor, the bottom part does not do that: it simply allows the student to refresh his/her (assumed) knowledge of electrical circuits, without reinforcing the metaphor and without providing any support for understanding magnetomotive force, other than

refreshing the analogy. The content of that page is also part of the abstract task of studying each unit of the LO; such a task prescribes to sequentially go through all the units, in the prescribed order. The order of the units is dictated by prerequisite constraints among concepts and the adopted teaching strategy.

2.2 Expression of a Learning Object

Expression of a message deals with the ways in which content can be perceived by humans. Therefore (except for brain-computer interfaces) expression requires that some *physical medium* is used to represent physical differences that we can sense (pressure waves in air, ink areas on paper absorbing light, pixels emitting light, etc.). While the physical medium depends on the interaction device being used, at a more abstract level we can identify *expression media*, which are usually classified into text, image, sound, or video. Components expressed in any of these media through specification of associated style attributes (f.e. font faces, colors, sizes, textures, voices, tones, frame-rate, etc.) are the building blocks of the expression of a LO, and constitute the substance of the expression. Appropriate media and style attributes are also used to represent controls upon which the student can act, f.e. the particular look and feel style used for representing the button to move to the next unit.

The form of expression involves space and time. Components (each with its own expression medium) can be combined together in a large variety of ways by choosing appropriate spatial, temporal, spatio-temporal or hyper structures. Text and images are static components that occupy some area in the bidimensional space provided by the browser window; their relative position is an expressive relationship. In other cases, text and images may be part of a dynamic medium (video): in the simulation that the LO of Figure 1 offers, the space occupied by images does not change over time, but other components and attributes do: text labels change between “5V” to “10V”, the picture of the knob changes when it is “turned”, the color of the lamp changes. In other LOs there might be sounds as well, with their own evolution over time, often synchronized with text, images or user actions.

The form of the expression is the result of choices that the LO authors made in order to achieve certain pragmatic effects on students; f.e. if an image is below some text, then it is either semantically nested or it follows the text. In both cases the student is expected to read the text first, which is more important.

In addition to temporal and spatial structures, within a LO there is also a hyper structure. Over time, depending on user-initiated events, the interface “moves” across this hyper structure. In general, the hyper structure reflects some of the relationships that were identified in the content and implements appropriate access and bearing information (f.e. navigation bars, drop-down menus, pop-up menus and tabs are four possible different expressions of the same combination of destination, location and access information). In our example, the hyper structure implements the linear structure prescribed by the task model, and is a sequence of units. Within a unit, several possibilities exist. For example, in the unit shown in Figure 1, the student can repeatedly run and re-run the simulation by switching the voltage knob; but the student cannot turn the current off once the circuit is on, unless the unit is restarted (by returning to the previous one and then moving on to the next). To run the simulation for the first time, the student has to click on the central area of the “on/off”

push-button; subsequently the only available action is clicking on the central area of the voltage knob. Consider also that the hyper structure may be time-dependent: f.e. until the student runs a simulation of the circuit there is no way to move forward to the next unit.

While in this discussion I focused primarily on a single LO, the same kind of properties apply to the LMS. Navigation bars, menus to invoke functionalities that are unrelated to the LO are all examples of other expression components that implement content that is part of the LMS.

LOs exploit multimedia (which is meant as *integration* of different media so that their combination achieves effects that are more than just the sum of the individual components) in order to obtain pragmatic effects on students, like enhancing attractiveness, achieving persuasion, capturing and retaining attention, increasing usability, enhancing memorability (f.e. it is known that images are more memorable than text, which is more memorable than spoken words, and that static images are more memorable than videos). However, unwanted interactions between media can produce also interference which prevents proper interpretation (f.e. when changes of text size rearrange the components in the page, perhaps removing some of the visual clues that are needed to interpret the page).

Multi-modality is not directly bound to the expression being used in the LO, as it depends on the kind of physical actions that students can execute (on appropriate physical interaction devices like keys, switches, joystick, knobs, wheels, cameras, etc.) in order to carry out common operations (like activating a button, selecting an option, dragging and dropping something, etc.).

3 Accessible Learning Objects

With "accessible LO" I mean a LO that is as effective for disabled students as it is for non-disabled ones [11], where "effective" is meant in terms of task accuracy and completeness and not in terms of teaching effectiveness. "Disabled student" include the typical categories of disabled users of the Web, namely: people with no vision using screen/Braille readers and keyboard; low-vision users of screen magnifiers; deaf users using sign language or non-audio information; motor disabled people using only the keyboard, special keyboards or mice, scanning devices, body gestures, or speech recognition; cognitively disabled people that may use speech recognition, word predictors when typing and may need to use simple language and additional icons.

Accessibility clearly depends on the user platform, which includes the operating system, its user interface components (f.e. KDE), the underlying accessibility architecture (f.e. MSAA), and the assistive technology (f.e. the screen reader, the screen magnifier, a modified keyboard, an on-screen keyboard, etc.).

Provided that the LMS is accessible, and that the user platform includes assistive technology required by the student, the expression of the LO depends on the user platform and on specific options set by the student during usage of the LO. For example, text could be read aloud; images could be "transformed" into spoken words; scenes of a video could be enriched with textual captions describing them; audio content could be transformed into textual transcripts. These changes in expression involve *inter-media* transformations (f.e. text to spoken words), *intra-media*

transformations (f.e. by changing the geometric properties of space when using a screen magnifier to enlarge the screen, or when changing the text size), *temporal* transformations through new synchronization of events (f.e. by using audio signals to notify a user of a screen reader that a certain feedback message has appeared in a location that differs from the current focus of interaction) or slow-down of an animation/simulation and *de-contextualization* of information (f.e. when the user of a screen reader extracts and lists all the links in a page, so that each link is rendered out of its original context).

Some of these transformations affect interaction modality. For example, in an accessible LO new operations may be made available, like the ability to extract and scan links in a page, or page headings, or the ability to jump directly to the content of the page, or the ability to move back and forth through items of a list. Since interaction devices used by students are not constrained by the LO author, to carry out these operations students may use any devices: only the keyboard, only a sip-and-puff sensor, a virtual keyboard on the screen, voice commands, etc. Obviously, not every combination of expression and interaction devices makes sense: f.e. blind users of a screen reader don't usually use mice because they do not provide sensible feedback.

Provided that a LO is accessible, a first necessary condition is that its content is available to wide range of students using several platforms. The LO can support transformations of its expression so that meaning is not lost (i.e. neither interest nor bearing information is lost). A second necessary condition is operability of the LO so that each relevant operation (needed to supply functional information or to operate on access information) can be performed through appropriate actions (i.e. operations specified in the task model part of the content are not lost). We call these conditions *invariance of content* with respect to changes in expression and interaction modalities.

3.1 Agents Affecting These Transformations

In order for these transformations to occur so that invariance of content holds, a number of enabling conditions need to be met.

1) The user platform should support all required transformations, through standard assistive technology, browsers, plug-ins and accessibility architectures.
2) The technologies used by the LMS and the LOs (f.e. HTML, CSS, JavaScript, SMIL, Flash, PDF) should be supported by the user platform, and the technologies themselves should support the required transformations.
3) Authors of LOs need to provide the required redundant expression in the different media that can be used by students (f.e. textual descriptions of video scenes) and required specifications to support transformation of expression (f.e. synchronization constraints so that captions are rendered at the right time).

Thus, from the perspective of the LOs authors, accessibility requires them to clearly identify all the content units (the substance and form of the content, see Section 2.1) and make sure that 1) interest and bearing information can be transformed into all possible media that might be available in students' platforms, 2) that the transformation are complete (f.e. all bearing information is transformed) and 3) that operability is guaranteed (f.e. all functional information and access information can be operated upon by students).

3.2 Modeling Content

One way to achieve invariance of content is to make sure that authors of LOs can actually think about content of LOs independently from their possible expressions. Appropriate notations can support such a process: Figures 2 and 3 show examples of content models expressed through conceptual maps, and through UML activity diagrams. Conceptual maps (Figure 2) can describe a taxonomy/ontology and highlight relevant concepts and their relationships. In this way, authors can lay down the domain of interest to be covered and decide which concepts will be dealt-with by the LO.

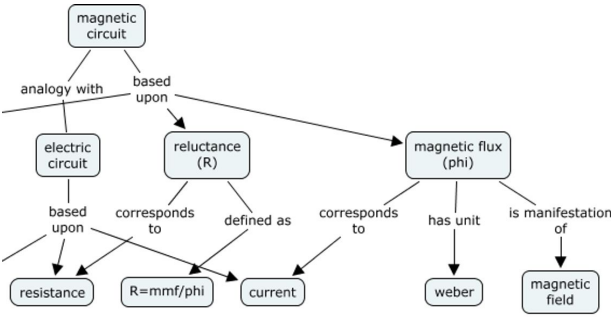


Fig. 2. Part of the basic domain concepts covered by the LO shown in Figure 1

Finally, Figure 3 shows part of an activity diagram [4] that specifies the activities that the LO does, interlaced with the activities that students are expected to perform. For example, the LO begins by presenting an information of interest stating the analogy between magnetic and electrical circuits. The task model specified by the diagram requires the student to act in order to move to the next unit, entitled with the location information "Magnetic force", and containing as information of interest an introduction to the concepts of electric force, voltage and current. A further student action leads to the description of the phenomena regarding the basic electrical variables, specified by the UML object called "electric circuit", another information of interest which, in the expression shown in Figure 1, is implemented through the simulation involving the images. The student can explore different scenarios (using the indexed access information that supports setting the voltage to 0, 5 or 10v) regarding these phenomena until s/he continues to the next unit.

An activity diagram is useful because it can highlight, as UML objects, the specific contents that the LO should present, synchronized with the activities these objects are associated with. For example, the object "electric circuit" is shown when the unit entitled "Magnetic force" is shown, but also when the students "activates" the subsequent system activity called "present details of elec. circ.". Such an object disappears afterwards.

What these diagrams represent is totally independent from expression. It could be implemented with the expressions shown in Figure 1 (namely text that changes, animated images, specific buttons, specific screens), or in a totally different way (like a voice narration of the electrical phenomena captured by object "Electric circuit").

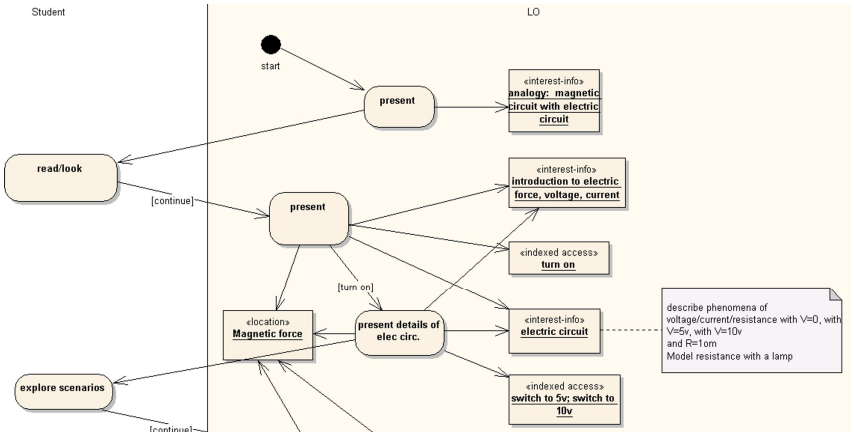


Fig. 3. Part of the activity diagram describing actions of the student and of the LO shown in Figure 1

The same freedom is available for implementing how the student is expected to move back and forth between units or how the different scenarios (0, 5 or 10v) are to be selected.

4 Discussion and Conclusions

Accessibility is a starting point for multimodal LOs, but unfortunately it is not enough. In fact, while accessibility guarantees levels of task effectiveness, specific expression choices made by the designer contribute to the learning effectiveness. As we've already seen, different media lead to different memorability degrees; interpretation of images requires less effort and time than text; spoken words might not be understood if there is a long interference with other sounds; static and moving images can be used to attract attention, they can provide affective clues, they can compensate text; sound icons can attract attention. Videos and animations can attract and hold students attention, but dominate on other media that are rendered at the same time. With proper use of expression, authors of LOs can implement contrast between components and can assign different relative importance to them by using emphasis. Unfortunately, automatic transformations of expression are not guarantee to preserve these properties.

As a consequence, content invariance does not guarantee the same level of learning effectiveness. In my view the only solution for having adaptable multimodal multimedia LOs (i.e. such that their expression change according to the use situation but without affecting the learning effectiveness) is to explicitly design different expressions, each geared towards a specific use situation [2]. Therefore, rather than relying on automatic adaptation of expression achieved through the transformations we discussed above, the author has to provide them explicitly. But also in such a case, content modeling is crucial, because authors can focus first on content and then, independently, design all the alternative expressions.

I believe that modeling content (separately from its possible expressions) is a fundamental step for enabling truly accessible and multimodal LOs. Appropriate content modeling notations and systems need to be adopted by designers, but fortunately they are already available. By separating content from expression, one can more easily determine which content units should be made available to students, how they could be made available to students, and how students might operate on them.

References

1. Brajnik, G., Toppano, E.: Creare Siti Web Multimediali, Addison-Wesley/Pearson Education Italy (2007)
2. Brajnik, G.: Achieving Universal Web Access through Specialized User Interfaces. In: Stary, C., Stephanidis, C. (eds.) User-Centered Interaction Paradigms for Universal Access in the Information Society. LNCS, vol. 3196, Springer, Heidelberg (2004)
3. Bocconi, S., Nack, F.M., Hardman, L.: Using rhetorical annotations to generating video documentaries. Technical report CWI INS E0501 (2005)
4. Conallen, J.: Building Web Applications with UML. Addison-Wesley, Reading, MA (2003)
5. Cooper, A., Reimann, R.: About Face 2.0: the Essentials of Interaction Design, Wiley and Sons (2003)
6. Hjemlslev, L.: Prolegomena to a Theory of Language, University of Wisconsin Press (1961)
7. Holzinger, A.: Fundamentals of Human-Computer Interaction (HCI) for e-Learning. In: Mittermeir, R.T. (ed.) Innovative Concepts for Teaching Informatics. Vienna (Austria): Ueberreuter C, pp. 157–159 (2005)
8. Instone, K.: Location, Path & Attribute Breadcrumbs (visited October 2006) http://www.user-experience.org/uefiles/breadcrumbs
9. Mann, W., Thompson, S.: Rhetorical Structure Theory: Towards a Functional Theory of Text Organization, Text, vol. 8(3) (1988)
10. Rosenfeld, L., Morville, P.: Information architecture, O'Reilly (2003)
11. Slatin, J., Rush, S.: Maximum accessibility: Making your Web Site More Usable for Everyone. Addison-Wesley, Reading, MA (2003)

Entertainment and Ambient: A New OLDES' View

M. Busuoli[1], T. Gallelli[2], M. Haluzík[3], V. Fabián[4], D. Novák[4], and O. Štěpánková[4]

[1] Ente per le Nuove tecnologie, l'Energia e l'Ambiente-ENEA, Brussels, Belgium
[2] Cup 2000 SpA, Bologna, Italy
[3] Charles University, Third Department of Medicine, Czech Republic
[4] Czech Technical University in Prague, Faculty of Electrical Engineering
massimo.busuoli@bruxelles.enea.it, teresa.gallelli@cup2000.it,
mhalu@lf1.cuni.cz, xnovakd1@labe.felk.cvut.cz,
fabiav1@fel.cvut.cz, step@labe.felk.cvut.cz

Abstract. This work introduces a new concept of supporting elderly at their homes. The whole framework is being developed under OLDES project: Older People's e-services at home. OLDES aims at developing a *very low cost and easy to use entertainment and health care platform* designed to ease the life of older people in their homes. The platform is based on a PC corresponding to *Negroponte's paradigm* of a € 100 device, giving the guarantee of an affordable system. OLDES provides: *user entertainment* services, through easy-to-access thematic channels and special interest forums supported by animators; and *health care facilities* based on established Internet and tele-care communication standards. As an example of OLDES platform implementation, two pilot projects are addressed: 100 clients pilot including 10 people with cardiological problems in Bologna, Italy and diabetes pilot in Prague, Czech Republic.

Keywords: Ambient intelligence, entertainment, user interface, diabetes and cardiac disease monitoring.

1 Introduction

Home health care service allows older persons to live as much as possible in their own ambient, with the members of their family, when clinical state and family consent that. Several projects, e.g. [3], [5], have studied and tested various approaches how ICT can improve this service addressed to not self-sufficient patients suffering with chronic diseases, to children with chronic diseases, to people suffering from serious handicap. It is integrated into the health network including hospital and territorial services to guarantee the continuity of the cure. The service is assured by a team of various operators: doctors, nurses, social workers that collaborate with the general practitioner. Sometimes it involves voluntary associations. Home health care service is carried out according to a personalized treatment plan, based on the assessment of the functional state of the patient and care problems oriented. Over three-quarters of the patients that utilize the home health care service is over 75 years old.

It is necessary to produce improvements as regards the actual home health care delivery [5] whether promoting healthy life style through an adequate scheduling of

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 511–519, 2007.
© Springer-Verlag Berlin Heidelberg 2007

thematic radio channels, or through the constant monitoring of vital parameters and health state, that allows a more adequate and well-timed treatment of possible diseases. Moreover opportunity of extending the service to a wider number of users and having a great number of available data, relevant to monitor and analyse the patient behaviour and clinical state, is not negligible.

1.1 Current Landscape

Up to date the community centres offering recreational and cultural services have been running by senior citizens themselves (volunteers) in full autonomy and open to all citizens, providing facilities such as bars and recreation centres, reading and TV rooms, organising exercise courses, leisure activities and social events. These centres are fully aware that, although the organisation and activities offered should be geared towards the current situation, they should also think ahead to future strategies regarding bringing people together and fighting isolation. The term **"isolation"** is used as it communicates the feeling of suffering, whereas "being alone" can be a choice.

The presented concept represents a step forward to use tools for exchange of information between humans on the web in order to reduce the loneliness and isolation of older people. In this context we aim to provide a complete lifestyle monitoring of the patient and to provide the huge (logistic, emotional and psychological) advantage to allow the old persons to stay home, in a less "hostile" and "unknown" environment than the hospital or health care housing. The low cost of the proposed device (around € 100) can be affordable by a Municipality like Bologna considering its advantages and the possibility for instance to save money by reducing, and in some case removing, long term institutionalisation of elderly people. So far local government has limited or rejected a large-scale adoption of new technologies because their high costs that were not affordable by the public funding resources. The scope of the project is to overcome this critical issue by developing a service accessible to the municipal budget and that can be delivered on a systematic basis to every old person of the town. The system can of course be extended to offer and manage many other services like booking services, transportation, etc.

1.2 Concept Vision

In order to explain how the project advances beyond the present state-of-the-art a brief representation of the vision is explained.

The goal of computer-aided gerontology is to maintain up-to-date information about the condition of thousand of aged people across a given region. This is the problem municipalities face today and that will increase dramatically. The current solution is to call them individually from a contact centre. This situation is critical: many European cities have more than thirty thousand of inhabitants aged 75 or more. An employee of a contact centre can contact at most 35 people a day. So, it is impossible to deal with such a situation. Moreover, if it is not possible to contact all these people, it is certainly not possible to gather them in a given place, so that they meet friends or participate to activities.

The suggested concept gives aged people living at home a digital companion, in the form of a low cost PC [2, 8, 10]. This device will implement various popular communication means, but in a way that makes them transparent and easy to use for the old and disabled people [11]. Without knowing it, they will be able to use the most recent communication technologies to communicate between each other, with their relatives and care services. In addition to this, their experiences will grow and improve the more they use the system. High quality experiences are characterised by the proactive and reactive targeting with highly relevant content, by the personalised proactive and reactive assistance and by the intelligent social networking features offered.

2 Method

The approach proposed is an integrated web-based system able to manage information and "knowledge" regarding the domestic environment and the biological conditions [9]. The system interfaces all home and vital devices by collecting information from automatic signals [1] and manages the entrainment channels. The technological system is composed of two parts: software and hardware.

2.1 Basic Concept

The main stakeholders are shown in Fig. 1. Different groups connected to the proposed system: thousands of elderly people (*clients*) and several animators and professionals. *Central system* refers to a place where the information about all the users is known. A direct connection between *doctor/social services* and elderly people can also take place but usually, it can be only done on the explicit request by the patient.

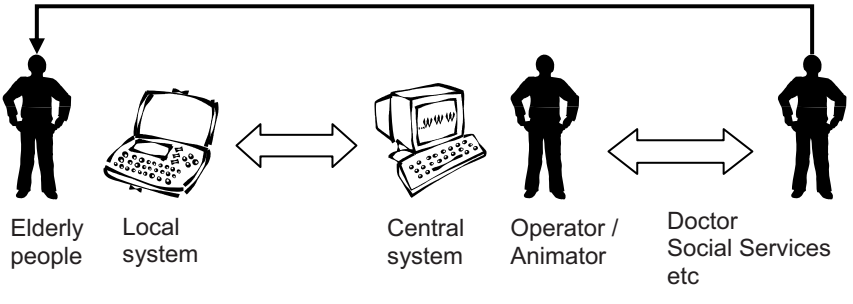


Fig. 1. Agents

Elderly people can get the needed feedback in various ways: through the bi-directional communication characterising the functioning of the entertainment services, by communication with the operators of the contact centre or directly with their GP. GPs (and/or clinicians) take part of the proposed system as well. Moreover,

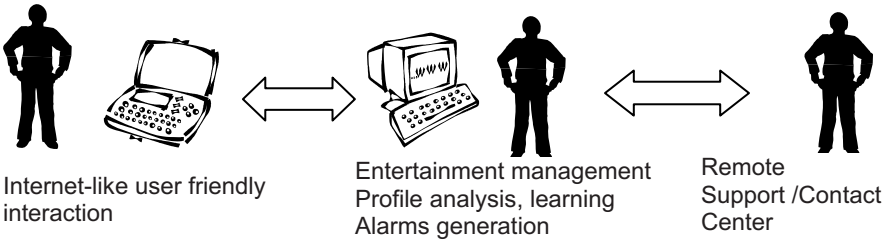


Fig. 2. Roles

the on-line analysis of the client's behaviour and health status allows health and social operators to give each client specific support if needed through the above-mentioned means.

The goal of the system is to provide both entertainment and communication services as shown in Fig. 2.

The **local system,** based on the low cost PC, runs the following services:

- Forums, thematic channels
- User-friendly feedback of the sensor data, health diagnostic, warnings or alarms.

The **central system** consists of:

- tools for the animators to manage the virtual communities;
- stores, retrieves and updates the user profile like a socio-sanitary dossier;
- tools for the management of the interactions with the elderly people and for the analysis of signals coming from home devices.

The **remote support** allows access for the doctors, social services, nurses, etc. This access can be given through

- a direct access to the central system;
- messages send by the central system, given the result of the analysis;
- or the animators.

2.2 SW Infrastructure

All developed SW is intended to be open source (Linux) and all HD used is low cost [7] with plug-in capability. The processing SW module consists of a decision support system using an integrated approach involving a rule-based expert system and machine learning based system. The goal of decision support is to supply the best recommendation under all circumstances.

2.3 Local Infrastructure

Central to the software is the "communication system" that facilitates communication between the user and Contact Centre concerning tele-accompany service, while monitoring system facilitates communication between Central Node and the installed domestic and personal devices and related sensors concerning social medical support.

The services are based on Internet technologies. The platform reuses existing established and new communication standards: e-mail, instant messaging, voice-over-ip as Skype - see Fig. 3.

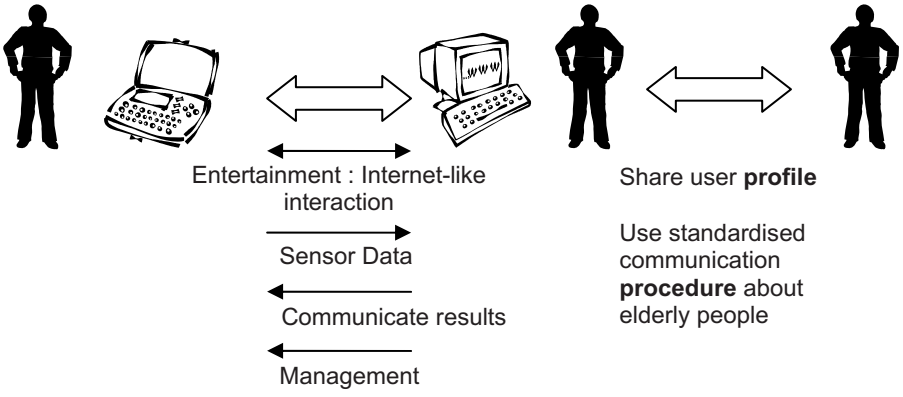


Fig. 3. Communications

The local sensors use a standard communication layer, such as BlueTooth, or Zigbee. Only one communication layer is used. It must be cost effective, low power and reliable – see Fig. 4.

2.4 Pilots Project

The system is being tested at two different locations: in Italy over a group of 100 elderly (including 10 suffering with cardio disease) and in Czech Republic over a group of 10 diabetic patients.

- Italy, where the user entertainment feature is extensively tested over a group of 100 elderly (in their homes) in conjunction with health monitoring (through specific sensors connected to the communication device) of a sub group of about 10 people with cardio problems.
- Czech Republic, where the communication feature is used in conjunction with the health monitoring (through specific connected to the communication device) to put in direct connection a hospital with a group of 10 diabetic patients staying at home.

These two pilots allow to prove the systems flexibility, easiness to use and transferability to different countries. The different approaches adopted in the two pilots test:

- the modularity of the system – the possibility to use it for different purposes with simple plug and play devices and sensors being everything managed centrally;
- its easiness to use – the users just have to wear the sensors or interact with the services via a simple user interface;

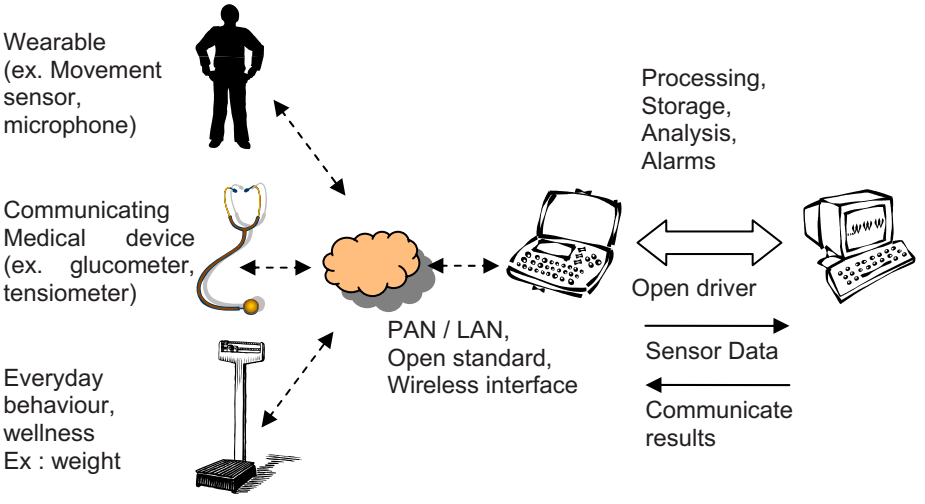


Fig. 4. Sensors

- its exploitation potential – the different characteristics of the two countries where the system is proven will provide the confidence (if the pilots will be successful) for the replication potential of the overall approach and technology.

2.5 Infrastructure Integration

The PC device is used to provide user entertainment (tele-accompany) and care facilities through a series of easy-to-access thematic channels. The PC mounts easy to use SW that can be utilised by the old person or his/her family [11]. The PC is integrated with multiple wearable and ambient sensors and devices that gather the information to be sent to a remote Central Node (CN).

The CN manages all the potential functions of the single local devices and facilitates communication both with the user and, as appropriate, with entertainer centres, remote call centre, relatives, social services providers, health care providers etc.

Sensors and devices consist of sensors for vital functions (as for instance ECG, respiration, skin and ambient temperature, heart rate, pulse oximetry SpO2, and diastolic and systolic blood pressure) monitoring and movement detection.

At the CN level, algorithms analyse the data received by the local PCs and manage automatic warnings to alert remote call centre if user behaviours and health parameters change suddenly (i.e. a person who has been not participating for a long time to e-conference sessions or a person having a very low skin temperature).

3 Results and Discussion

The Municipality of Bologna aims at offering the whole system to the elderly over 75 of the town (about 30.000 persons could be involved in the medium period) and at

setting up a centre for monitoring and analysing the needs and parameters of elderly in order to allow communication among different communities (elderly themselves, entertainer centres, health and care professionals). The municipality of Prague investigates the specific implementation of the project on group of diabetes patients.

3.1 Diabetes Case Study

Elderly population commonly suffers from multiple diseases/co-morbidities and requires treatment by a combination of different medicaments. This population therefore carries the most significant risk of acute worsening of the diseases themselves and/or side effects of the medicaments and their combination, respectively. It is estimated that more than 15% of population older than 65 years suffer from diabetes mellitus. Clustering of diabetes with arterial hypertension, obesity, dyslipidemia and other pathologies is commonly referred to as metabolic syndrome. The presence of this syndrome markedly increases the risk of cardiovascular morbidity and mortality of diabetic population.

Concerning diabetes pilot project, a preliminary study has been carried out using blood pressure monitoring of 250 patients aged over 60. The system consists of blood pressure monitor equipment, which automatically reports results of up-to-date patient's measurements to the doctor's monitoring and decision support system using wireless communication. The following sensors have been used: Finapres system, standard electronic tonometer OMRON M4I, classical mercury tonometry and electronic tonometer OSCIOO [6] measuring directly oscilometrical pulsations. Significant differences (even over 30%) in blood pressure measurements between classical tonometer based on mercury and electronic ones have been found. That is why further physiological signals as SpO_2 must be added to the final alerting system following thus the OLDES concept. The results will be compared to those obtained in another project considering the same diseases using different technologies [4].

3.2 OLDES' Benefits

The main innovation in the OLDES proposal is that by using a low cost platform for delivering interactive services and monitoring:

- the elderly maintain contact with the carer and each other via interactive services;
- the solution is low cost both in software and hardware terms because of the low cost PC and the ASP model;
- it is easy to use for the elderly: you just turn on the device to access the interactive services provided by the carer;
- the services can be personalised to the needs of the elderly. The animator knows the profiles and names of the people he is interacting with;
- the server keeps the history of measurements from sensors for analysis and raising health related alarms;
- measurements are made at the client's location and monitored centrally at the call centre. This allows the doctors to spend less time visiting patients who don't need it and provides better supervision of patients (each patient has a health agenda that is monitored by the call centre);

- the client technical infrastructure is very simple. The PC acts as a communication gateway and doesn't manage sensor data locally;
- the carer can monitor more patients that before because of the monitoring assistance and the raising of alarms by the monitoring software;
- the proposed platform supports a modular approach to services: services can evolve and be provided as needed without much work on the client side. It can be done centrally by the application service provider;
- it promotes social life for the elderly by supplying them with interactive services. Some of the services allow the isolated elderly to talk to each other;
- the approach is sustainable because of the low cost of the proposed platform. It is easy to add new services and new sensors;
- by monitoring patients you can avoid going to the hospital when it is not necessary;
- the platform makes some of the most popular internet services available in an easy to use form for the elderly;
- it supports mobility by allowing the elderly to be anywhere in their home and still interact with the contact centre thanks to the wireless communication. For example if the patient is sick he/she can still interact with the call centre as long as the wireless communication is working.

For the municipalities and their contact centres, the project provides a number of solutions for improving the efficiency and effectiveness of assistance and support processes dedicated to elderly and disabled people. Based on *automatically generated Content Intelligence* (what topics are consumed, which are the popular ones, which are missing / content gaps, emerging topics), *Customer Intelligence* (user demographics, user thematic preferences, user social relationships, user behavioural patterns, user health status) and *Context*, the system generates relevant messages and alarms and trigger the appropriate workflow(s) / process(es). This dramatically improves the scalability of the present systems, decrease personnel costs (operational and schematisation) and enhance the quality of the services. With such a system, the municipalities, health-care service providers and other related service-providers are able to better know their target audiences and to offer them better services at lower costs.

4 Conclusion

To sum up, at the end of the project there will be an easy to use, plug and play system with a different costing level based on the profiling of the person to which it will be provided:

- a base level (offered to all) which includes the communication and tele-accompany and which is simply based on the low cost PC and the open source SW with a target price of around € 100 per person.
- an intermediate level with the addition of simple sensors (like for example the ambient temperature) for the management of generic monitoring situations (i.e. hot summer periods management)
- an upper and tailored level with an addition of health monitoring sensors depending on the health profile of the user.

Acknowledgments. The project was supported by STREP IST-2005-2.6.2-AAL OLDES (Older People's e-services at home).

References

1. Cook, D., Das, S.: Smart Environments: Technology, Protocols and Application. Wiley-Interscience (November 2004)
2. Cammaerts, B., Van Audenhove, L., Nulens, G., Pauwels, C.: Beyond the digital divide: reducing exclusion, fostering inclusion. VUB Brussels University Press (2004)
3. CONFIDENT http://www.etw.org/2003/case_studies/soc_inc_confident.htm
4. Doc@home http://fp6uk.ost.gov.uk/Page.aspx?SP=2034
5. eInclusion@EU http://www.einclusion-eu.org/default.asp?MenuID=8
6. Fabián, V., Fejtová, M.: Telemedical system for monitoring of blood pressure. Advances in Electrical and Electronic Engineering, pp. 213–215 (2005)
7. ink-media.com
8. Kubitschke, L., Stroetmann, V.N., Hüsing, T., Stroetmann, K.A.: Older People and Information Society Technology: Results from the European SeniorWatch Survey. In: ICAT – International Conference on Assitive SeniorWatch (2002), published by British Computer Society (2002)
9. Nussbaum, G., Miesenberger, K.: The Assistive Home – More than Just Another Approach to Independent Living. In: ICCHP 2004. LNCS, vol. 3118, pp. 891–897. Springer, Heidelberg (2004)
10. http://www.seniorwatch.de
11. Stephanidis, C.: User Interfaces for All - Concepts, Methods, and Tools. Lawrence Erlbaum Associates, Mahwah, NJ (2001)

Virtual and Mixed Reality Interfaces for e-Training: Examples of Applications in Light Aircraft Maintenance

Johannes Christian[1], Horst Krieger[2], Andreas Holzinger[3], and Reinhold Behringer[1]

[1] Innovation North, Faculty of Information and Technology
Leeds Metropolitan University, Leeds, UK
r.behringer@leedsmet.ac.uk,
johannes.christian@gmail.com
[2] ipcenter.at, Vienna
horst.krieger@ipcenter.at
[3] Institute of Software Technology and Interactive Systems (IFS)
Vienna University of Technology, Vienna, Austria
holzinger@ifs.tuwien.ac.at

Abstract. There is evidence that recent developments in Augmented Reality (AR) technology has the potential to be applied as pervasive media on multiple devices in different ways and contexts, especially with low-cost devices including Mobile Augmented Reality (MAR) applications on smart phones or Pocket-PCs. In this paper we present a framework in order to combine the pervasive e-education concept with augmented reality content for e-training. We analyze current research, discuss some examples from ultralight / light sport aircraft maintenance and show how to apply this framework generically. We present a learning engine to deliver this special type of content and provide a further outlook of future research. A user-centered approach must ensure that the developments can stimulate motivation and enhance performance of the end users in different training sessions. The main benefit is, that the end users are enabled to better perceive complex, technical facts, systems and components.

Keywords: e-training, augmented reality, pervasive e-education, performance support, human computer interaction, usability engineering.

1 Introduction and Motivation

Recent developments [1-7] have shown that Augmented Reality (AR) technology has the potential to be applied as a pervasive medium on a variety of platforms in different ways, ranging from Head Mounted Displays (HMDs) to hand-held Mobile Augmented Reality (MAR) applications on smart phones or Pocket-PCs [2],[3],[4], [5],[7]. AR, which links real and virtual worlds, opens new paradigms for interacting with computers and enables new ways of experiencing learning [1],[6] These new paradigms emerged – on the one hand – from the need of modern businesses for implementing a process of continuous innovation, permanent construction of knowledge and transfer of knowledge to their workers [8]. On the other hand, companies are more and more forced to offer context-based, time-of-need-based

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 520–529, 2007.
© Springer-Verlag Berlin Heidelberg 2007

learning and performance support tools to support their workers and customers in fulfilling their job duties [9].

In the specific context of an aviation application scenario, it is interesting to note that not only pilots but also aviation maintenance personnel is a highly mobile international end-user group. While performing operational and maintenance tasks on the flight line or at different maintenance facilities, directly on the airplane or in the workshop, these people need the possibility to work independent of location, while still following the rigid instructions for continuing airworthiness.

This paper explores educational uses of Virtual Reality (VR) and AR technology especially for application workflows in training and maintenance of ultralight / light sport aviation. It describes the status of an ongoing project, highlighting the initial studies and the proposed outline of the system to be developed. State-of-the-art research was analyzed, with focus on the technical challenges, recent developments, new interaction and interface opportunities and a description of possibilities to provide virtual multimedia content for learning with electronic means for aircraft maintenance.

Given the necessary technology, psychological research becomes both possible and essential in order to provide a user-centered development and to ensure the best possible assistance for the end-users. Results must always be incorporated at systemic level, which in this context includes – for example – the consideration that the proper combination of learning content and learning context is of critical importance. Learning content must be developed in such a manner that meaningful and sustainable learning effects with the learners can be created [10]. At this point, it should be noted that there are also research results indicating that there is no influence of media on learning results [11]. However, Driscoll and Carliner [12] argue that well designed learning content can engage learners by making them active participants in real-world problem-solving, thus creating meaningful and sustainable learning experiences. We suppose that this is especially true for VR and AR applications, because of their use of immediate interactive engagement.

Practical examples will demonstrate the learning benefit in aircraft operation, maintenance and installation. Visualization, animation and simulation content applied in VR and AR open a wide range of possible application scenarios and can change static content into dynamic virtual multimedia content [13]. These are composed of 3D objects, images, illustrations, text, sound, 3D avatars and videos.

1.1 The Specific Case of Aircraft Maintenance

Several research studies in this field fail to combine knowledge from x-learning [5] with AR/VR technologies and their effect on how these can improve the way we interact with learning material. We strive for overcoming this deficiency. For example, in combining the Magic Book concept [6] and the magnifying glass approach [16], we demonstrate within a prototype scenario the way in which our perception of and interaction with content in standard training literature could alter so as to create meaningful and sustainable learning experiences and job aids supporting pilots, maintenance technician and employees' tasks.

Sometimes small problems occur during the maintenance process, which the maintenance technician would be able to solve with the relevant on the job support

having access to Learning on Demand (LoD) instructions and to trouble shooting or diagnosis references on his current available device (just-in-time learning). In such performance support cases, we need to consider models that detail the person / time / place / context / device / medium in the relevant setting [19] and thereby provide a form of quasi-intelligent assistance. This concept is based on the assumption that jobs are conglomerates of little tasks [20] with each task being now supported by a job aid in the form of a LoD sequence.

Furthermore, it is also of importance for aviation maintenance engineers to participate in continuous training and mandatory recurrent training sessions. Consequently, this research will also provide examples of important learning support to supplement traditional face-to-face (F2F) and hands-on/on the job training such as carrying your virtual instruction with you on your laptop/smartphone and showing the procedures to be performed in a multimedia augmented environment.

2 Background and Related Work

Our work on AR/VR based e-training in aviation is built on related work in (handheld) AR/VR, pervasive e-education and the combination of both. In this section we briefly review and highlight past work in these areas.

There are already applications of AR for training and education in mathematical content [21], biology, mechatronics, engineering education [20] and aeronautical engineering [3]. Further on some examples show the application of Mixed Reality (MR) for museums [23], [1], [24] and military use [25]. Another growing field of applications can be found in edutainment [26] and gaming [27].

In the past years, a trend has been observed, bringing AR into mass market using handheld devices (mobile phones) [7] and more integrated head-worn displays (HMDs) [28]. Authoring toolkits for AR-applications have matured in their usability and are close to become easier in handling of their User Interface (UI) e.g. similar to the authoring of text or illustration in technical communication. In our experiments and research, we currently use the AMIRE toolkit [29] to perform the various test settings on a laptop. AMIRE already was tested successful in an industrial scenario in an oil refinery [30]. The extension has been developed AMIRE-ES [2] for use with mobile devices. A similar AR programming and authoring platform is “Studierstube” [31]. It presents a framework for the development of mobile, collaborative and ubiquitous AR applications. Another practicable solution for AR on low-end handheld devices in a markerless environment is presented and a topic for research within the ULTRA-Project [32].

Adobe Acrobat 3D provides the possibility to create PDF-documents with embedded 3D models and animations. With the freely available current version of Adobe Reader even quite complex 3D models can be viewed and manipulated on a standard PC.

Pervasive E-Education [14],[15],[16] has potential to be applied for Technical Training in corporate environments in different business domains when presented as a concept to assist teaching and learning for Life Long Learning (LLL) and Continuing Engineering Education (CEE). However, so far nobody explored to combine the concept of pervasive e-education and the integration of AR and 3D learning content in different settings. Our work addresses this gap.

3 Issues in Ultralight/Light Sport Aviation Training

Based on the above framework, our goal is to set up training scenarios for an international target group in ultralight / light sport aviation (LSA), which ranges from an aircraft operator (pilot) to the aircraft maintenance technician, who performs maintenance direct on the airplane or repair in his workshop. According to their specific requirements they have different tasks, technical documentation and instructions for continued airworthiness they need to be trained on. An aircraft operator (pilot) needs to be trained on how an aircraft is operated correctly, e.g. how to start the engine and which limits not to exceed. The aircraft maintenance technician (AMT, A&P) needs more to know when, where and how to perform maintenance checks, e.g. to inspect the spark plug condition and gap on a piston aircraft engine. An AMT in the workshop has to have specific training on e.g. how to repair a component like a gearbox. Also in case of problems the aircraft operator (pilot) and/or the AMT also may need to get support to perform troubleshooting, e.g. where to start with search for cause of failures.

Currently all this training is mainly done in a traditional classroom setting for getting the theoretical background. The practical training part is mainly performed as a hands-on training in the workshop direct on the aircraft or aircraft engine. During this time, the highly qualified person is away from the workplace or has to travel to a certain place in the world to get this specialised training. Therefore, it is obvious that training can benefit from AR/VR based e-training in various ways as the pilot or technician can access training resources via the world-wide-web from everywhere. For the practical training aircraft parts and aircraft systems can already be viewed in advance in a virtual world on the workplace or supporting information can be given by AR technology in standard training material/ technical documentation or direct on the real (training) object in a concurrent way. But in order to gain a wide accepted basis for such a framework social, technological as well as psychological aspects have to be considered and the stakeholder requirements need to be carefully analysed.

4 System Architecture and Implementation

In this section we describe the components of our system, which is under way. Most of the components are in the process of implementation.

4.1 The Learning Engine

Learning Management Systems (LMS) or course / content management systems (CMS) are quite common in e-learning environments providing content. For a virtual training centre an LMS can integrate administrative, curricular, communicative and social infrastructure. This central platform can provide various kinds of content ranging from text-based up to high-tech interactive, multimedia and 3D. As a structured repository the system provides capacity to administrate AR content for download to the different end-devices. The Open Source CMS “Moodle” [33] is very popular and used over a wide range. It offers a variety of UI languages for an international community of training participants. We have chosen Moodle as our

prototype-platform because it shows to be quite flexible and a good choice for a first research prototype of learning engine. The central learning engine will administrate, store and provide the relevant training content available as text, multimedia, 3D or AR presentation. The user can access the virtual training center (learning engine) from different end-devices like laptop, MDA by being connected via LAN, WLAN, 3G/UMTS or HSDPA to the internet and access if authorized the different learning and performance support material right at the time when needed. The overall architecture can be seen in figure 1.

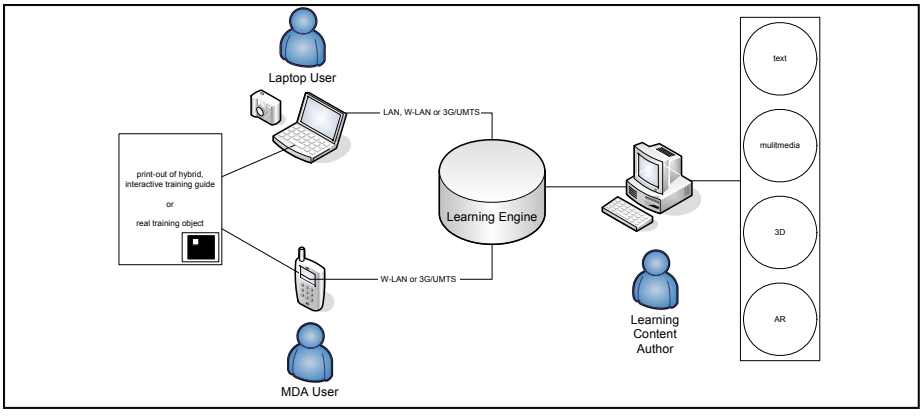


Fig. 1. The architecture of our learning environment

4.2 Content Typology

In the content authoring process content can be added for each type of training along the Reality-Virtuality Training Continuum [34] (figure 2) based on the original mixed

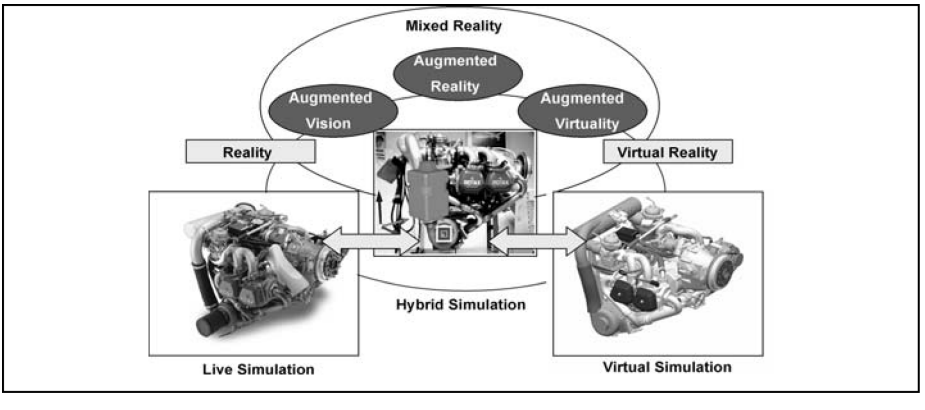


Fig. 2. Reality-Virtuality Training Continuum demonstrated in our specific application on an aircraft engine. Left: real aircraft engine. Middle: real aircraft engine blended with a virtual element of an overflow bottle (cooling system). Right: virtual engine as designed in 3D.

reality continuum [35]. Thus, we can establish what we call a content typology. In figure 2 we show that starting from the left in reality in the classroom the technician gets normally theoretical instructions and training directly on the object like in our example on the aircraft engine. This process is assisted by training aids like cut-away parts, single parts or by training manuals showing technical drawings and illustrations. Further from the left to the right the augmentation increases until a user gets a virtual aircraft engine designed in 3D by CAD specialists presented. Between these 2 worlds, different levels and kinds of augmentation by e.g. video, text, 3D objects, sound can assist the AR user.

4.2.1 Standard Training Manuals

These training manuals are edited in a standard desktop publishing process together with Subject Matter Experts (SMEs) or as an extract from the technical documentation. These are provided in hardcopy or distributed electronically in a PDF.

4.2.2 3D Knowledge Objects (3KO)

As per definition of NGRAIN Corporation [34] Knowledge Objects and 3D Knowledge Objects (3KO) are defined as self-contained, reusable chunks of data. The extension to 3D Knowledge Objects (3KO) means that they contain a 3D simulation of a piece of equipment. The concept of 3KOs combined with recently available Acrobat 3D forms a quite immersive and interactive way to reuse for e-training 3D data of the Digital Product (e.g. modelled and available form Pro Engineer (Pro/E)) created in the Digital Engineering Process. A 3D enriched PDF file allows to interact with the 3D models with a standard mouse and offers different functionalities like rotate, pan, zoom, spin, set different lighting, play animation and much more.

4.2.3 Augmented Reality Knowledge Objects (ARKO)

Extending the definition of Knowledge Objects for Augmented Reality (ARKO) means, that they include the methods for use in an AR environment. Digital Parts (imported e.g. from Pro/E into blender for conversion or animation and then authored in AMIRE) are reused, but this time blending virtuality and the real world. The user can interact with the virtual object e.g. using a web-camera or MDA (figure 4) as a tangible interface. In moving the interface or the print-out of the hybrid, interactive

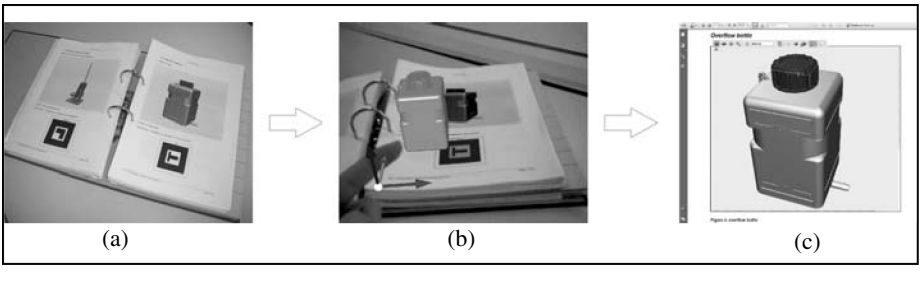


Fig. 3. Reality-Virtuality Continuum for a hybrid, interactive training guide: (a) the traditional paper-based material; (b) AR view into the training guide; (c) interact with 3D training content

Fig. 4. Augmented Reality view on a MDA prototyped by through a Hybrid Mock-up (real MDA with transparent slide (paper) mock-up of 3D model)

training guide or the physical training object (in our case tracked by a marker card) a new way of learning is possible. Depending on the choice of user and the availability of the different media and tools learning in different ways is possible (figure 3).

4.3 Hardware/Software

For the technological experiments following hardware and software was chosen and showed to be technically sufficient:

a) Hardware:
A standard laptop (with internet access) resolution ranging from 800 x 600 px to 2048 x 1536 px with external 1.3 Megapixel Web Camera and a MDA Vario II with Windows Mobile 5, Quadband, HSDPA, WLAN and USB, TFT screen with a resolution of 240 x 320 px, internal 2 Megapixel digital camera.
b) Software:
We selected the following Open Source software packages: Moodle Server (CMS), MySQL (database), PHP (scripting language), PYHTON (dynamic object-oriented programming language), Blender (software for 3D modeling, animation, rendering etc.), Blender Pocket (port of Blender for Pocket PC), AMIRE (framework for Authoring MR under LGPL), Open Office (multiplatform and multilingual office suite), Mozilla Firefox (web-browser). Additionally, the following commercial software was used: ProEngineer (parametric feature-based 3D Solid modelling CAD software), Adobe Acrobat 3D (assembling 3D training guide PDF and convert PRO/E files), Acrobat Reader 8.0 (PDF reader), ClearVue PDF (mobile PDF reader), Remote Display Control (in order to access the MDA), Microsoft Active Sync (in order to synchronize PDA with PC), Microsoft Internet Explorer Mobile.

4.4 User Interaction Combining Learning Engine with Learning Content

The user accesses the portal by opening a web browser and typing the URL of the virtual training centre. Depending on his access rights the focus either stays in the public area or enters the secured training area. On the platform there is a possibility to choose between the different content e.g. opening the PDF file on the screen or printing it. Using the PDF file allows to open the 3D data of a digital product/part. For

use on a PDA, this data must be split into an adopted PDF content and into a separate 3D content, as PDF viewer on Windows Mobile does not offer yet 3D viewing capability. In the current implementation, the AR content must be downloaded locally to the laptop or the MDA and be opened there with an appropriate AR viewer. As easy authoring/prototyping for handheld AR content we have set up a rapid prototyping method to be able to get a look and feel for handheld AR in using transparent slide mock-ups.

5 Conclusion and Future Work

In this paper, we showed that combining the x-learning concept with AR and VR has potential for future pervasive e-training in aircraft operation and maintenance as compared to the traditional mainly paper-based F2F-training methods. Research allows us to presume that this approach is the first example of this concept and technology with the target audience in aviation. The respective target groups are provided with different types of information access and content and may choose between various types of content representation. While reading PDF files or print-out material, our mobile approach enriches the traditional way of training through the embedded VR content, which allows information to be viewed in 3D. AR allows the 3D training material to be examined at the workplace; a workbench in the workshop or the use of a training device in a real environment subject to the weather. However, much further research is necessary in order to bridge the gap between the potential benefits of these new technologies and the usability of these systems. We will concentrate on developing interfaces based on user centered design principles, particularly in the area of authoring tools.

Further experiments will be based on our first findings and will encompass different AR devices including lightweight HMDs and Ultra-Mobile-PCs (UMPCs). Augmented Reality 2.0 aims at enabling people to work together in order to create and share AR content online and distribute it for AR applications. Another possibility is the collaboration with AR avatar on the chair next to them in order to improve their learning performance. For this purpose the AR based learning paradigm must be evaluated against learning from print based material and classic desktop VR based content.

References

1. Abawi, D.F., Dörner, R., Haller, M.: Efficient Mixed Reality - Training Applications Development. In: Proceedings of Training, Education & Simulation International (TESI 2005) Maastricht, NL. pp. 1–10 (2005)
2. Haller, M., Stauder, E., Zauner, J.: AMIRE-ES: Authoring Mixed Reality once, run it anywhere. In: Proceedings of the 11th International Conference on Human-Computer Interaction (HCII) 2005 Las Vegas, USA (2005)
3. Haritos, T., Macchiarella, N.D.: A Mobile Application of Augmented Reality for Aerospace Maintenance Training. In: 24th Digital Avionics Systems Conference, Avionics in a Changing Market Place: Safe and Secure,: Washington, DC (2005)

4. Henrysson, A., Ollila, M.: UMAR Ubiquitous Mobile Augmented Reality. In: Proceedings of the MUM 2004 Maryland (2004)
5. Makri, A., et al.: ULTRA: Light Augmented Reality Mobile System. In: Proceedings of the ISMAR 2005 Vienna (2005)
6. Pan, Z., et al.: Virtual reality and mixed reality for virtual learning environments. Computers & Graphics 30(1), 20–28 (2006)
7. Wagner, D., Schmalstieg, D.: Handheld Augmented Reality Display. In: Proceedings of IEEE Virtual Reality (VR 2006) IEEE: Alexandria, Virginia, USA. p. 321 (2006)
8. Takeuchi, H., Nonaka, I.: Hitotsubashi on Knowledge Management. John Wiley & Sons (Asia) Pte Ltd, Singapore (2004)
9. Bielawski, L., Metcalf, D.: Blended eLearning. Integrating Knowledge, Performance Support, and Online Learning. HRD Press Inc, Amherst (2003)
10. Allen, M.: Michael Allen's Guide to e-Learning. Building Interactive, Fun, and Effective Learning Programs for Any Company, Hoboken: John Wiley & Sons Inc. (2003)
11. Clark, R.E.: Media Will Never Influence Learning. ETR & D 42(2), 21–29 (1994)
12. Driscoll, M., Carliner, S.: Advanced Web-Based Training Strategies. Pfeiffer / John Wiley & Sons, Inc, San Francisco (2005)
13. Liarokapis, F., et al.: Multimedia Augmented Reality Interface for E-learning (MARIE). World Transactions on Engineering and Technology Education 1(2), 173–176 (2002)
14. Holzinger, A., Nischelwitzer, A.K., Kickmeier-Rust, M.D.: Pervasive E-Education supports Life Long Learning: Some Examples of X-Media Learning Objects (XLO). In: 10th IACEE World Conference on Continuing Engineering Education: Vienna, Austria (2006)
15. Holzinger, A.: Pervasive E-Education und Life Long Learning [February 11, 2007] (2006) http://www.competencesite.de/wissensmanagement.nsf/0/80d3291373725948c12570f8002b7455
16. Holzinger, A., Nischelwitzer, A., Meisenberger, M.: Lifelong-Learning Support by M-learning: Example Scenarios. ACM eLearn Magazine, vol. 5 (2005)
17. Billinghurst, M., Kato, H., Poupyrev, I.: The MagicBook - Moving Seamlessly between Reality and Virtuality. IEEE Computer Graphics 21(3), 6–8 (2001)
18. Rekimoto, J.: The Magnifying Glass Approach to Augmented Reality Systems. In: Proceedings of the International Conference on Artificial Reality and Tele-Existence '95 / Conference on Virtual Reality Software and Technology (ICAT/VRST '95) Tokyo, Japan. pp. 123–132 (1995)
19. Quinn, C.N.: Delivering the Dream- Models for Intelligent Assistance, OtterSurf Laboratories (2004)
20. Hale, J.A.: Performance-Based Management: What Every Manager should Do to Get Results, San Francisco: Pfeiffer, An Imprint of Wiley (2004)
21. Kaufmann, H., Schmalstieg, D.: Designing Immersive Virtual Reality for Geometry Education. In: Proceedings of IEEE Virtual Reality Conference 2006 (VR 2006) Computer Society: Alexandria, Virginia, USA. p. 7 (2006)
22. Liarokapis, F., et al.: Web3D and Augmented Reality to support Engineering Education. World Transactions on Engineering and Technology Education 3(1), 11–14 (2004)
23. Liarokapis, F., White, M.: Augmented Reality Techniques for Museum Environments. Mediterranean Journal of Computers and Networks 1(2), 95–102 (2005)
24. Connect The CONNECT Project [Internet] [Accessed 11 February 11, 2007] (2004) Available from:http://www.connect-project.net
25. Hughes, C.E., et al.: Mixed Reality in Education, Entertainment and Training: An Interdisciplinary Approach. Computer Graphics and Applications 26(6), 24–30 (2005)

26. Wagner, D., Schmalstieg, D., Billinghurst, M.: Handheld AR for Collaborative Edutainment. In: Proceedings of 16th International Conference on Artificial Reality and Telexistence (ICAT), pp. 85–96. IEEE, Hangzhou, China (2006)
27. Liarokapis, F.: An exploration from virtual to augmented reality gaming. Virtual Reality Simulation 37(4), 507–533 (2006)
28. Grobman, A.: Lumus Video Eyeglasses [Internet], Israel, Lumus Ltd. Available from: [Accessed February 11, 2007] (2007) http://www.lumus-optical.com
29. AMIRE The AMIRE Project (IST-2001-34024) [Internet] [Accessed February 11, 2007] (2002) Available from: http://www.amire.net
30. Haller, M.: Mixed Reality @ Education. In: MApEC - Multimedia Applications in Education Conference, FH JOANNEUM / Lisa Zimmermann: Graz. pp. 12–18 (2004)
31. Schmalstieg, D., et al.: The Studierstube Augmented Reality Project Presence: Teleoperators and Virtual Environments 11(1), 33–54 (2002)
32. Riess, P., Stricker, D.: AR on-demand: a practicable solution for augmented reality on low-end handheld devices. In: AR/VR Workshop of the Germany Computer Science Society. Coblence, Germany
33. Dougiamas, M., Taylor, P.C.: Moodle: Using Learning Communities to Create an Open Source Course Management System. In: Proceedings of the World Conference on Educational Multimedia, Hypermedia and Telecommunications (EDMEDIA 2003) Honolulu, Hawaii. pp. 171–178 (2003)
34. Kirkley, S.E., Kirkley, J.R.: Creating Next Generation Blended Learning Environments Using Mixed Reality, Video Games and Simulations. TechTrends 49(3), 42–53 (2005)
35. Milgram, P., Kishino, F.: A taxonomy of mixed reality visual displays. IEICE Transactions on Information Systems 77(12), 1321–1329 (1994)
36. NGRAIN, An Introduction to 3D Knowledge Objects, NGRAIN Corporation: Vancouver, Canada (2004)

A Holistic Approach to the Evaluation of E-Learning Systems

M.F. Costabile, T. Roselli, R. Lanzilotti, C. Ardito, and V. Rossano

Dipartimento di Informatica, Università degli Studi di Bari, 70125 Bari, Italy
{costabile,roselli,lanzilotti,ardito,rossano}@di.uniba.it

Abstract. This paper describes the eLSE methodology to evaluate e-learning systems. By combing a specific inspection technique with user-testing, eLSE allows inspectors, even not having a wide experience in evaluating e-learning systems, to perform accurate evaluations. The inspection technique is based on the use of evaluation patterns which precisely describe the activities to be performed during inspection. For this reason, it is called Pattern-Based (PB) inspection. An experiment has shown that novice evaluators are able to come out with good results, confirming the efficiency and the effectiveness of PB inspection.

Keywords: Usability inspection, e-learning system, controlled experiment.

1 Introduction

Life-long education is addressed from many points of view in the Information Society. E-learning systems exploit Artificial Intelligence techniques to better fit the needs of their users. A common objective of researchers' work in this field is to develop e-learning systems of high quality. One of the main goals of an learning system is to avoid any distraction and to keep the whole content fresh in the learners' minds. In the specific case of e-learning, the challenge is to create an interactive system that does not confuse learners. We often find that an e-learning system is a mere electronic transposition of traditional material, presented through rigid interaction schemes and awkward interfaces. When learners complain about Web-based training or express a preference for classroom-based instruction, it is often not the training, but rather the confusing menus, unclear buttons, or illogical links that scare them off.

The user interface of an e-learning system can become a barrier to effective learning and information retention: if it is not well designed, the user can feel lost, confused or frustrated [6]. Moreover, technology should not become a barrier. Users with different hardware and software equipment should be able to exploit the e-learning artefacts, possibly through customized access procedures.

Different techniques and methodologies are reported in literature that aim at ensuring the quality of e-learning systems. However, the identified solutions appear as a short blanket, able to cover only some of the multiple aspects that characterize the complexity of the problem [16]. Moreover, the proposed criteria are often only

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 530–538, 2007.
© Springer-Verlag Berlin Heidelberg 2007

vaguely stated, so that any actual measurement is left to subjective interpretation. Some authors have proposed that the usability heuristics summarized by Nielsen [10] and Shneiderman [14] can be applied to evaluate e-learning systems interfaces [5], [13]. Other researchers assert that usability testing needs additional consideration in the light of the e-learning environments currently available, and propose a list of heuristics adapted to such a context [11], [12], [15].

This paper illustrates the *eLSE* (e-Learning Systematic Evaluation) methodology, which aims at increasing the reliability and the effectiveness of e-learning evaluation by proposing a structured and systematic approach to it [8]. eLSE methodology systematically combines inspection with user-testing. The main novelty of this methodology is the use of evaluation patterns which precisely describe the activities to be performed during the inspection. It is then called *Pattern-Based (PB) inspection*. Evaluation patterns precisely describe which objects of the application to look for, and which actions to perform during the inspection in order to analyse such objects. In other words, they describe the tasks the evaluators must perform during their inspection activity. For this reason, in previous work this technique was also called Abstract Task (AT) inspection [2]. By exploiting evaluation patterns, even less experienced evaluators are able to come out with more complete and precise results. In order to perform a more systematic evaluation, the proposed approach focuses on two separate aspects of an e-learning system: the platform and the educational modules. A controlled experiment has shown a promising advantage of PB inspection over user-testing and heuristics evaluation, demonstrating that evaluation patterns, specifically defined for e-learning systems, are efficient tools to drive evaluators and improve their performance [2].

The paper has the following organization. Section 2 describes the eLSE methodology and shows how the eLSE methodology suggests to perform the evaluation process. Section 3 describe the empirical evaluation performed to validate the PB inspection. Finally, Section 4 concludes the paper.

2 eLSE Methodology

eLSE (e-Learning Systematic Evaluation) methodology prescribes a structured flow of activities. The main idea of eLSE is that the most reliable evaluation results can be achieved by systematically combining inspection with user-based evaluation. Several studies have outlined how two such methods are complementary and can be effectively coupled to obtain a reliable evaluation process [10]. Accordingly with those studies, eLSE suggests to couple inspection activities and user-testing, and precisely indicates how to combine them to make evaluation more reliable and still cost-effective. Each evaluation process starts by having evaluators inspecting the application and identifying possible problems and troubles. User testing is then conducted, whenever necessary, to validate some inspection findings with real users. The eLSE methodology suggests to analyze an application along the specific dimensions (i.e. hypermedial technology, presentation, user activity, educational process, application proactivity) that address the appropriateness of design with respect to the peculiar nature and purposes of the e-learning systems [8]. Finally,

eLSE proposes an inspection technique based on the use of evaluation pattern, that are specifically defined for e-learning systems. For this reason, it is called *Pattern-Based (PB) inspection*. The evaluation patterns precisely describe which objects of the application to look for and which actions the evaluator must perform in order to analyze such objects.

2.1 eLSE Activity Flow

eLSE methodology organizes the evaluation process activities into a preparatory phase and an execution phase. The preparatory phase aims at creating a conceptual framework that will be used to carry out actual evaluations. The output of the preparatory phase can be easily shared among different evaluators, or different evaluation laboratories, that have similar interests and evaluate such applications from similar points of view. The preparatory phase consists of the definition of a library of evaluation patterns. The execution phase is performed every time a specific application must be evaluated. It mainly consists of inspection, performed by evaluators. If needed, inspection can be followed by user testing sessions, involving real users. At the end of each evaluation session, the evaluators must provide designers and developers with a proper evaluation feedback. The activities of the two phases are described in the following sections.

The Preparatory Phase. In the preparatory phase, the definition of a specific set of evaluation patterns must be carried out. By considering the literature on e-learning systems, the experience of human-computer interaction experts, and the results of observing real users interacting with e-learning systems [3], evaluation patterns, that address specific characteristics of e-learning systems to support the evaluators in their inspections, have been defined.

Evaluation patterns guide the evaluator's activities by describing which elements of the application to look for and which actions the evaluator must perform in order to analyse such elements. In this way, even novice evaluators, with lack of expertise in usability and/or application domain, are able to come out with more complete and precise results. The evaluation patterns are defined with the intent of capturing the evaluation expertise of few expert inspectors within our research group, who are able to analyze e-learning systems. Trying to explain our method of analysis, we have developed a number of rules about what to do and what to look for, while performing an evaluation. This is how the evaluation patterns started to shape up capturing the actions of expert inspectors, and trying to explain them to inexperienced people. Essentially, the evaluation patterns make possible to maximize the reuse of evaluators' know-how, by capturing usability inspection expertise, and by expressing it in a precise and understandable form, so that it can be easily reproduced, communicated, and exploited. They therefore allow evaluators "… taking advantage of any of the efforts done in previous works, to reduce the effort needed to achieve a new one" [9]. The concept of pattern, as a cognitive tool to capture human expertise and to make it re-usable, was introduced by the architect C. Alexander, applied to architecture and urban planning: "…Each pattern describes a problem which occurs over and over again in our environment, and then describes the core of the solution to that problem in such a way that you could use this solution a million times over…" [1].

The evaluation patterns are precisely formulated by means of a template that provides a consistent format and includes the following items:

- *Classification Code and Title*: they univocally identify the evaluation pattern, and succinctly convey its essence.
- *Focus of Action*: it shortly describes the context, or focus, of the evaluation pattern, by listing the application components which inspectors have to evaluate.
- *Intent*: it describes the problem addressed by the evaluation pattern and its rationale, trying to make clear which is the specific goal to be achieved through the use of the evaluation pattern.
- *Activity Description*: it describes in detail the activities to be performed during the use of the evaluation pattern.
- *Output*: it describes the output of the fragment of the inspection the evaluation pattern refers to.

During the PB inspection, evaluators apply the evaluation patterns, performing the detailed activities stated in them.

The list of evaluation patterns provides a systematic guidance to the evaluator on how to inspect an application. Most evaluators are very good in analysing certain features of interactive applications; however, they often neglect some other features, strictly dependent on the specific application category. Exploiting a set of evaluation patterns ready for use allows evaluators with limited experience in a particular domain to perform a more accurate evaluation.

In comparison with a classic inspection, such as heuristic evaluation, the use of evaluation patterns during the inspection provides several advantages: 1) they are defined by observing expert inspectors that performed various inspections, thus capturing evaluators' expertise; 2) they provide information about application domain, tasks and users, since this information is provided in the activity description item; 3) they enforce standardization and uniformity, by suggesting a language that inspectors will use for writing their evaluation report; 4) they reduce time and cost needed for an accurate inspection, as demonstrated by controlled experiments in which the PB inspection has been compared with other evaluation techniques, namely heuristic evaluation and user observation [2], [4].

Referring specifically to point 2 above, it is worth noting that evaluation patterns convey necessary information to the evaluators about application domain, tasks and users. In the heuristic evaluation, to get this information, evaluators must undergo some form of training.

The PB inspection for e-learning systems proposes an important distinction between e-learning platform (the container) and educational modules (the content). The e-learning platform is the interface environment which may offer a number of integrated tools and services for teaching, learning, communicating, and managing learning material. The educational modules, also called Learning Objects (LOs), are the specific pedagogical content provided through the platform. Evaluation patterns for platform have been defined; they differ from the evaluation patterns defined for e-learning modules, since different features and criteria need to be considered.

Two evaluation pattern categories have been identified [7]:

Quality in use analyzes the technological and structural characteristics of e-learning systems. Both in the case of e-learning platform and learning modules, the quality in use patterns consider the technological infrastructure in terms of ease of use, communicational effectiveness, functions with respect to the designed educational goals and activities, and ease of use for learners.

Content learnability considers the educational characteristics of e-learning systems. In the case of the e-learning platforms, the content learnability patterns refers to the capacity of the platform to allow learners to learn the presented contents. The content learnability category refers to the learning module capacity to transfer the course content to the learners and to make it understandable in a satisfactory way.

For both categories, quality in use and content learnability, evaluation patterns for the platform and evaluation patterns for the LO have been defined.

Table 1 reports an example of evaluation pattern for e-learning platform in the category "Content learnability"; the title is "Ease of use of the authoring tools".

Table 1. Example of evaluation pattern for evaluating e-learning platform

P_CL_05: EASE OF USE OF AUTHORING TOOLS
Focus of action: authoring tools
Intent: evaluate the introduction modalities of new educational material
Activity description:
– Use authoring tools to define new contents
– Try to insert a new document by testing all commands available
– Try to update a document that has been just inserted
Output: a description reporting if:
– There is not mechanisms that permit the creation of document in standard formats (AICC, IMS, SCORM)
– It is difficult to update the material just inserted.

Table 2 reports an example of evaluation pattern for educational module in the category "Content learnability"; the title is "Appropriateness of the language".

Table 2. Example of evaluation pattern for evaluating learning objects

LO_CL_12: APPROPRIATENESS OF THE LANGUAGE
Focus of action: educational content
Intent: to verify the language used in the text of a selected topic
Activity Description:
– Choose a topic
– Analyze content from a language point of view
– Simulate inability to understand a term and try to look for its definition
Output: a description reporting if:
– The language used is not appropriate to the target user
– Difficult terms are not defined in the text (before their use) or in a glossary.

The Execution Phase. Execution phase activities are carried out every time an e-learning system must be evaluated. They include two major jobs: a *systematic inspection* and a *user-based evaluation*.

Systematic inspection is performed by evaluators. During the inspection, the evaluator uses the evaluation patterns to perform a rigorous and systematic analysis and produces a report in which the discovered problems are described, as suggested in the output item of the used evaluation patterns. The list of evaluation patterns provides a systematic guidance to the evaluator on how to inspect an application. Most evaluators are very good in analysing certain features of interactive applications; however, they often neglect some other features, strictly dependent on the specific application category. Exploiting a set of evaluation patterns ready for use allows evaluators with limited experience in a particular domain to perform a more accurate evaluation.

User-based evaluation is conducted only when there is disagreement among the evaluators on some inspection findings, so that validation with real users becomes necessary. The activity description of the evaluation patterns is a formulisation of the user tasks, thus it is the starting point to immediately formulate experimental tasks which can guide users in critical situations encountered by the evaluators during inspection. During the evaluation, a sample of users is observed while they are performing experimental tasks and relevant data are collected (users' actions, users' errors, time for executing tasks, etc.). The outcome is a collection of raw data, which are coded and organized in a synthetic manner in the result summary, and then analyzed.

The last activity of the execution phase aims at providing the designers and developers of the application with an organised evaluation feedback. The result of this activity is an evaluation report describing the problems detected, possibly revised in the light of the user testing outcome, using the terminology provided in the evaluation patterns for referring to system objects or interface elements, and for describing critical incidents. This standardised language increases the precision of the report and decreases the risk of misunderstandings.

3 Empirical Validation of the PB Inspection

An empirical validation of the PB inspection technique has been carried out [2]. The study involved seventy-three senior students of a Human-Computer Interaction (HCI) class at the University of Bari in Italy. They were divided in three groups that were asked to evaluate a commercial e-learning system by applying the PB inspection, or the traditional heuristic evaluation, or a thinking aloud technique. Each group was assigned to one of the three experimental conditions. The heuristic inspection group had to perform an heuristic evaluation exploiting the "learning with software" heuristics [15]. In the user testing group, every evaluator observed a user during the interaction with the e-learning system using the thinking aloud technique. Twenty-five students of another computer science class use the e-learning system, thus acting as users for the user testing. Finally, the PB inspection group used the inspection technique based on evaluation patterns, specifically defined to evaluate e-learning systems.

A week before the experiment, all participants were given a 1-hour demonstration of the application to be evaluated. A few summary indications about the application content and the main functions were introduced, without providing too many details. A couple of days before the experiment, a training session of about one hour introduced participants to the conceptual tools to be used during the experiment. Each group took part in a specific training session.

Data were collected in a group setting, but each participant worked individually. The study consisted of two experimental sessions lasting three hours each. During the first session, participants evaluated the e-learning system by applying the technique they were assigned. Members of the three groups had to find usability problems in the application, and to record them on a booklet. One day later, each evaluator was asked to type the problems they discovered on an electronic form. This was done to prevent readability problems during data analysis. In the electronic booklet, each participant reported the description of the problem, where it occurred, and how it was found. Finally, the evaluator attributed a severity rating to the problem on the scale from 1 (I do not agree that this is a usability problem at all) to 5 (usability catastrophe).

3.1 Results

Two expert usability evaluators independently examined all the electronic forms in order to identify single and unique usability problems. This analysis led to the identification of 247 problems and 49 non problems, or statements which reported not understandable content or unverifiable information. Non-problems accounted for 7% of the statements written by the inspectors applying the heuristic inspection, and 3% of the problems written by the inspectors applying the PB inspection or performing user testing.

On average, participants reported significantly more problems applying the PB inspection (mean = 21), than performing user testing (mean = 9.62) or the heuristic inspection (mean = 12.22). The difference is significant as demonstrated by the results of an Anova with evaluation technique as between-subjects factor ($F(2,70) = 25.45$, $p < .001$). Post-hoc tests applying the Bonferroni's correction indicated that the effect was due to the higher number of problems reported in the PB inspection condition, while there was no significant difference between heuristic inspection and user testing.

An overall efficiency index was computed, which reflected the number of problems each participant found in 10 minutes. The Anova indicated that overall efficiency is affected by experimental condition, $F(2,70) = 3.80$, $p < .05$. Post Hoc comparisons (Bonferroni's correction) showed that PB inspection was the most efficient technique (mean = 1.19 problems in 10 minutes, standard error = .08). There were no differences between heuristics and user testing, in the order mean = .91, standard error = .08 and mean = .90, standard error = .08.

The study also revealed that different techniques addressed different type of problems. User testing and heuristic inspection helped to highlight problems common to all interactive systems, whereas the PB inspection focused also on specific problems of e-learning. These findings confirm the peculiar characteristics of PB inspection that specifically addresses the e-learning domain, provided that appropriate evaluation patterns are defined.

4 Conclusions

Various evaluation methodologies and techniques can be considered and applied for evaluating e-learning systems. The paper has discussed issues related to the evaluation of this particular class of applications. In particular, we have proposed an evaluation methodology, called eLSE (e-Learning Systematic Evaluation), that prescribes a structured flow of activities. eLSE suggests that reliable evaluation can be achieved by systematically combining inspection with user-based evaluation and it precisely indicates how to combine them to make evaluation more reliable and still cost-effective [8].

eLSE proposes an inspection technique aiming at allowing inspectors, possibly not having a wide experience in evaluating e-learning systems, to perform accurate evaluations. It is based on the use of evaluation patterns which precisely describe the activities to be performed during inspection. For this reason, it is called Pattern-based (PB) inspection.

The advantage of the PB inspection over other evaluation techniques has been demonstrated by a controlled experiment. Seventy-three participants were divided in three groups that were asked to evaluate a commercial e-learning system by applying the PB inspection, or the traditional heuristic evaluation, or a thinking aloud technique. The experiment results have shown that evaluation patterns, defined for e-learning systems, are effective and efficient tools to drive evaluators and to improve their performances [2].

During the experiment, the students expressed positive opinions on the use of evaluation patterns. They said that the guidance offered by the evaluation patterns allows them to perform the inspection more accurately. Students perceived to have more control on the inspection process and, consequently, they were confident to obtain good results. Furthermore, the two experts that have examined the electronic forms filled by the inspectors noticed that patterns enforce standardization and uniformity of inspection results, since all inspectors are guided by the output item writing their evaluation report. Moreover, the PB inspection is capable to address specific issues of e-learning better than other techniques exploited in the experiment.

Acknowledgments. The financial support of Italian MIUR through VICE project is acknowledged.

References

1. Alexander, C., Ishikawa, S., Silverstein, M., Jackobson, M., Fiksdhal-King, I., Angel, S.: A Pattern Language. Oxford University Press, New York, USA (1977)
2. Ardito, C., Costabile, M.F., De Angeli, A., Lanzilotti, R.: Systematic Evaluation of e-Learning Systems: an Experimental Validation. In: Proc. of NordiCHI 2006 Oslo, Norway (October 14-18, 2006) pp. 195–202 (2006)
3. Ardito, C., Costabile, M.F., De Marsico, M., Lanzilotti, R., Levialdi, S., Roselli, T., Rossano, V.: An Approach to Usability Evaluation of e-Learning Applications. International Journal Universal Access in the Information Society 4(3), 270–283 (2006)

4. De Angeli, A., Matera, M., Costabile, M.F., Garzotto, F., Paolini, P.: On the Advantages of a Systematic Inspection for Evaluating Hypermedia Usability. International Journal of Human-Computer Interaction, Lawrence Erlbaum Associates, Inc 15(3), 315–335 (2003)
5. Dringus, L.: An Iterative Usability Evaluation Procedure for Interactive Online Courses. Journal of Interactive Instruction Development 7(4), 10–14 (1995)
6. Kruse, K.: Web rules: effective user interface design (2000) Available at http://www.learningcircuits.org/may200/may2000_webrules.html
7. Lanzilotti, R.: A Holistic Approach to Designing and Evaluating e-Learning Quality: Usability and Educational Effectiveness, PhD dissertation, Dip. Informatica, Università di Bari, Bari, Italy (2006)
8. Lanzilotti, R., Costabile, M., Ardito, C., De Angeli, A.: eLSE Methodology: a Systematic Approach to the e-Learning Systems Evaluation. Educational Technology and Society 9(4), 42–53 (2006)
9. Nanard, M., Nanard, J., Kahn, P.: Pushing reuse in hypermedia design: Golden rules, design patterns and constructive templates. In: Proc. ACM HT, pp. 11–20 (1998)
10. Nielsen, J.: Usability Engineering. Academic Press, Cambridge, MA (1993)
11. Notess, M.: Usability, User Experience, and Learner Experience. (2001) Available at http://www.elearnmag.org
12. Quinn, C.N., Alem, L., Eklund, J.: A pragmatic evaluation methodology for an assessment of learning effectiveness in instructional systems. (1997) Available at http://www.testingcentre.com/jeklund/Interact.htm
13. Schwier, R.A., Misanchunk, E.R.: Interactive Multimedia Instruction. Educational Technology Publications, Englewood Cliffs, NJ (1993)
14. Shneiderman, B., Plaisant, C.: Designing User Interface. Addison-Wesley, Washington, D.C. (2004)
15. Squires, D., Preece, J.: Predicting Quality in Educational Software: Evaluating for Learning, Usability, and the Synergy between them. Interacting with Computers, vol. 11(5), pp. 467–483 Elsevier Science Ltd, Great Britain (1999)
16. Zaharias, P., Vasslopoulou, K., Poulymenakou, A.: Designing On-Line Learning Courses: Implications for Usability (retrieved December 2006) (2002) from http://www.japit.org/zaharias_etal02.pdf

How Can HCI Factors Improve Accessibility of m-Learning for Persons with Special Needs?

Matjaž Debevc, Mateja Verlič, Primož Kosec, and Zoran Stjepanovič

University of Maribor, Smetanova 17, 2000 Maribor, Slovenia
{matjaz.debevc,mateja.verlic,pkosec,stjepanovic}@uni-mb.si

Abstract. Mobile learning, or m-learning, respectively, can be stated as an upgrade of e-learning, valuable for mobile users using new mobile technology and telecommunication systems. The article presents the influence of m-learning system intended for supporting a remote engineering application. Here, we examined the usability of the application designed for people with special needs taking into account the main human-computer interaction factors. User interface for mobile systems served as a platform for examining the functionalities of a micro-controlled system. Two usability testing methods, involving some persons with special needs, were conducted in order to examine the usability of the system. The results gained are valuable for designing improved, future-oriented m-learning applications.

Keywords: m-learning, accessibility, usability, evaluation, persons with special needs.

1 Introduction

The impact of computer usage and telecommunication media development increasingly influences the education. In the framework of this evolution, e-learning was developed, which enables professors and students, together with information and communication technology (ICT), to access necessary materials or to perform informing, distance communication and evaluation [1]. Successful introduction of technologies, like Bluetooth, GSM (Global System for Mobile Communications), WAP (Wireless Application Protocol, UMTS (Universal Mobile Telecommunications System) and since 2006, UMTS networks in many countries have been or are in the process of being upgraded with High Speed Downlink Packet Access (HSDPA), sometimes known as 3.5G [2]. All this mobile technology, which offers better data transfer from day to day, leads to transfer from e-learning to mobile learning (m-learning).

M-learning as modern educational field offers an option of capturing course content and other materials over the mobile networks and enables communication between tutors and students and among students internally. In this case, mobile tools like mobile phones and recently also Personal Digital Assistants (PDA) with phone function are used as accessories.

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 539–548, 2007.
© Springer-Verlag Berlin Heidelberg 2007

Researches have confirmed that learning using ICT can be easier and more effective. [3] in is his research confirms, that students spent 70% less time studying in certain cases when using ICT. This was especially valid for materials, where animations and video presentations were available. Moreover, Chou and Lou [4] and Schulz et al. [1] confirmed in their research, that distance learning is effective and useful.

Due to this effectiveness of education using ICT, the question raises, whether m-learning can be equally suitable and effective, and particularly whether we can use it for education of persons with special needs. And if so, how should the outlook and usability of these systems be in the future? The fact is, that people with special needs are the population with their own demands and problems while working and learning.

To review the impact of m-learning systems on the education effectiveness, we created an example of m-learning application for the technical environment (remote engineering) and controlled the usability of the application on people with special needs, taking into account human-computer interaction factors. This paper presents following items:

- The basics for design of accessible e-learning materials according to the analysis of needs and requirements of the users with special needs.
- Explanation of the chosen design of m-learning application and working principles of the system. In this purpose, we have created miniature model of the apartment and named it m-house. The model consists of micro-controlled system for remote control with use of handheld and mobile connections to Internet.
- Presentation of experiences gained while testing the m-learning application with people with special needs.

2 Basic Accessibility Guidelines for Designing and Structuring Accessible m-Learning Applications

World Wide Web, which can be used for m-learning, has its biggest power in universality and wideness. Therefore accessibility factor is becoming one of the most important planning and educational application's development aspects for all, no matter on their impairment. Accessibility is the field that helps ensuring that all the users can read and use interactive mobile applications.

One of the ways to successfully create user interfaces is to design the accessible material using the Web Content Accessibility Guidelines (WCAG) from Web Accessibility Initiative (WAI) [5]. These guidelines are planned to be used generally in web site design; however these guidelines differ to some extend when designing and structuring m-learning applications for people with special needs.

To determine guidelines for designing and structuring accessible m-learning applications we have run a user analysis using questionnaires, brainstorming and feedback reporting.

Table 1 presents some of the most important guidelines for structuring and designing accessible e-content, which were determined as result of the analysis considering the WAI guidelines.

Table 1. Basic guidelines for designing m-learning applications for persons with special needs

Impairment	Solution for m-learning content development
Partial sight	• Larger images • Using large number of tabs to allow larger images • Font size increment option • Color contrast control • Additional description of images • Increased number of audio information • Audio based navigation • Hardware based accessibility (screen magnifiers)
Blindness	• audio description of images and video • audible subtitles, • audio based navigation, • audio text description • vibrating option for PDAs • Hardware based accessibility (screen readers, additional Braille keyboards for PDA's)
Hard of hearing	• Subtitles • Stronger image and video oriented content • Voice recognition for feature selection • Option of textual communication with tutor
Deaf	• Sign language video, • video subtitles, • additional glossary in sign language, • easy reading content
Impaired dexterity (such as arthritis, tremor, lack of limbs)	• Hardware based accessibility (additional keyboards)
Impaired cognition (such as memory, concentration)	• Logical and easy navigation, • very easy reading content, • larger number of images

Additional guideline that has been common to all the groups, was a request for automatic recognition of user's position in application's environment, so that only relevant features are offered to the user.

It is interesting that there does not exist m-learning application for educational purpose, which would include all required accessibility methods in a single system. It is a common approach, that e-learning material is developed for each specific population independently. This means, that many systems for various users are being developed instead of single system for all the users.

3 Design Guidelines for m-Learning Considering HCI Factors and Tools

Today we are not only the information consumers, but also information producers – the use of e-learning, distant learning and mobile learning is rapidly growing. Therefore it is even more important to consider diversity of potential users that might be using our e-materials. Many users are unable to view our advanced graphics and video because of visual impairments, hear audio because of hearing impairments, orient themselves in a cluttered page because of learning disabilities, or select small web elements because of mobility impairments. Introducing technology to learning means more independence, participation and productivity of people with disabilities. However, technology itself is not enough if it is not used in an accessible way. There have been many attempts to increase the accessibility of digital resources and services. Different organizations, agencies and consortia are trying to improve the current situation, one of them being the World Wide Web consortium (W3C) [6]. Designing an m-learning system for people with disabilities can be quite challenging due to disability specifics.

We have based our initial planning and development of m-learning system on the roadmap defined by different studies and resulting usability guidelines, [7, 8, 9]. However, we have also set our own guidelines divided into following phases:

- *Designing a system and evaluation*

  Before designing a system it is necessary to analyze and evaluate existing e-materials to identify possible accessibility issues. The current research suggests the best way to start is with collecting different design solutions proposed by different people. This is so called parallel design [7]. Afterwards an iterative design approach is taken which consist of creating various paper or computer prototypes that can be tested by target users. The changes of the system are based on feedbacks from these users. This iterative process of testing and making changes continues until usability goals, defined for the system we are developing, are met [10].

  The next step is to provide appropriate content – content that is first relevant and then adapted to the requirements and abilities of the people with different kinds of impairments. For most of these users this means short, concise text with as many explanations of unknown terms as possible.

  When the relevant content is chosen it is crucial to focus on performance before deciding on the aesthetics of design. This is even more important for people with special needs because they show substantially lower levels of system response time delay acceptance. It is also important to set usability goals that include success rates and the time it takes users to find specific information. One of such goals is the number of mouse clicks needed for a person with special needs to access desired part of the materials.

  Several automatic evaluation methods exist that can be applied to conduct initial evaluations of the website. Automatic evaluation is the evaluation performed by software that can recognize certain types of design problems. There are many such tools available on the internet such as A-Prompt [11], Bobby [12], Truwex online check [13] and many more. Although automatic

evaluation methods can be very useful, they cannot replace evaluations with typical users.

- *Optimizing the user experiences*

  In this phase the main concern is to facilitate and encourage effective human-computer interaction by applying different HCI factors. Some of the most important factors are: (i) the information displayed should be in directly usable format; (ii) the organization of the content should be highly intuitive and logical; (iii) a user should not wait more than few seconds for a page to load; (iv) an appropriate assistance an feedback need to be provided whenever user has to wait; (v) designers should avoid unsolicited popup widows; and (vi) users task sequences should be consistent with how users typically do their work. Users can benefit from all these factors if they are applied during design.

- *Accessibility of e-learning materials*

  The accessibility of e-learning materials can be ensured by following several HCI accessibility standards and guidelines such as US Federal Accessibility Standards (Section 508) [14], World Wide Web Consortium Web Accessibility Initiative (WAI) [15], ISO [16], ITU [17] in other national legislations. Finding and applying relevant Assistive Technologies is also of great importance for the accessibility.

- *Computer hardware/software and design of web pages*

  Within the constraints of various types of resources (time, money), it is usually almost impossible to design for all users. Before making final decisions on design it is recommended to identify the hardware and software used by people with disabilities. The best way is to design, develop and test for the most common browsers and to consider different operating systems and different internet connection speeds. The internet connection speed is of special importance when (streaming) video is used for deaf and hard of hearing persons. Several Web Accessibility Evaluation Tools [18] can be used for the evaluation.

- *Web page design*

  Web page design is the mostly exposed part of the design process, where it is essential to pay attention to HCI factors affecting it. The principles for visible language should define arrangement of different information blocks and graphics on the screen, the typography, the information flow, visual identity, animation, colors and structures. Web page design should be directed by organization, springiness and communication [19].

- *Navigation*

  Navigation refers to the methods used to explore the information within a website. Navigation is vital for people with special needs, because for them is even more important to know their current location on the webpage and how they can return to the beginning or how to navigate to a certain point in the materials. Tables of contents with active links to relevant chapters in materials are especially useful if materials are long and structurally complex. Glossary of unknown terms with additional possibility of showing the explanation in the form of a video is supremely beneficial for deaf and hard of hearing persons.

Although accessibility issues and efforts are primarily concerned with people with special, it is good to know that access and accessibility of web sites, e-learning and m-learning systems can make learning easier for ALL people--not just disabled individuals.

4 Practical Example of Usage of m-Learning in Technical Environment

Having in mind the guidelines for designing and implementing of m-learning systems listed in previous section, we have decided to test the possibility of using m-learning for persons with disabilities.

We have designed and developed a system, named m-house. The system consists of wooden model of a house (Fig. 1), which includes micro-controlled system for control and inspection of devices, connected to Internet. User interfaces for a PDA were implemented in LabView [20]. Using appropriately adapted user interface, devices, which are integrated in the house model (e.g. lights, windows, doors, ventilation) can be remotely controlled via mobile access to Internet.

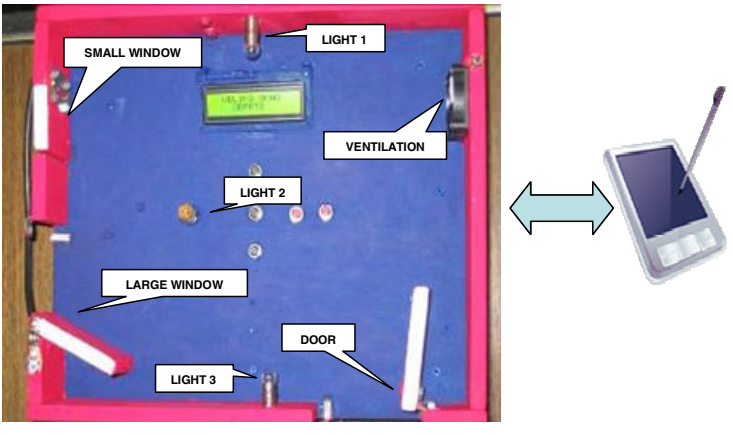


Fig. 1. M-house model (miniature model of an apartment with active elements that can be controlled via PDA)

The students, including persons with special needs, use this application to get acquainted with ways to use micro-controlled systems, control methods for distance control systems and methods to design user interfaces.

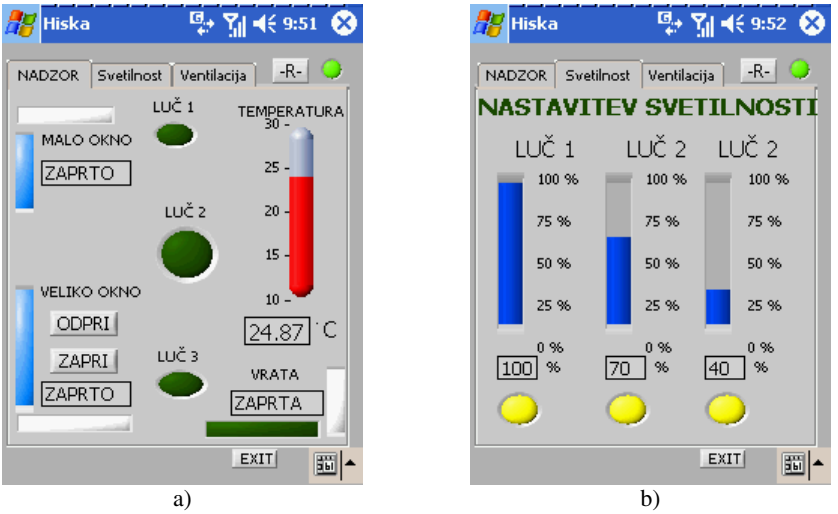
Other educational values of the system are:

- Acquaintance with new area of mobile technologies (not only telecommunication but also application use)
- acquaintance with methods for wireless data transfer (wireless, GPRS, UMTS)

- acquaintance with non-standard use of interfaces (people are used to use interfaces for personal computers, while PDA's smaller display requires the student to become aware of new demands, on how to solve certain visual limitations)
- Learning to use application wherever and whenever. PC application can be used only behind the table, where PC is placed. PDA or mobile application can be used wherever we are
- learning to develop application development in software for mobile technology
- Learning how to create less memory consuming applications. PDA's working memory is limited to few MBs, while application for PC can be of size of few GBs. Therefore it is necessary to learn to minimize the size of application and use only the necessary elements
- Acquaintance with data transfer savings. Data transfer using mobile applications is usually paid by amount of data transferred; therefore a good knowledge of communication protocol and knowledge of data flow minimization is necessary.

The user interface for mobile systems (handheld in our case) serves as platform for insight into the apartment and function of micro-controlled system. The interface was design with LabView software and offers several different tabs to control and inspect the elements. Figure 2 shows two different parts of same user interface (developed in Slovenian language) – inspection and light system control interface.

The development has been driven by the request for larger images for visually impaired people. This can be easily seen on figure 2, where each element is larger as it had been developed in the first phase.



a) b)

Fig. 2. User interface of m-house model. a) Elements' control b) Illuminating power settings.

This user interface gives us the insight on the status of windows and doors (open/closed), status of ventilation, lights (on/off) and the level of illuminating power. With virtual pressing of the light button, we can change its binary state (on/off). Changing the status of the ruler on side changes the level of illuminating power (between 20% and 100%)- User interface includes the temperature indicator, where current temperature value (measured every second) is shown graphically and with the number.

5 Usability Testing

For usability testing we conducted an experiment, involving people with special needs. Since the number of participants was limited only to three we chose methodologies, which gather qualitative information.

We combined two methods for gaining valuable qualitative data: the thinking-aloud method and the unstructured interview. During the test, the participants were asked to verbalize their thoughts, feelings and opinions while interacting with the system. They needed to perform a set of tasks while they were using the application. This method allows evaluators to understand how the participants approach the interface and what considerations they keep in mind when using the interface. So the main benefit with this method is a better understanding of the participant's mental model and interaction with the system [21]. After participant finished their tasks, the evaluator performed an unstructured interview. In this method, evaluators formulate questions about the application based on the kind of issues of interest. After that they interview the users to ask them these questions in order to gather information desired. For the optimal analysis of tests, both methods were audio taped.

As we mentioned before we tested three users. Two candidates were partial of sight; one with 95% of blindness, the other with 88%. The third candidate was hard of hearing with deafness of 97% per Fowler scale.

In the first stage, users performed the thinking aloud method, where they had to accomplish ten tasks, which were carefully selected by the evaluator. The first user (95%) had to use a magnifying glass, because the contrast and the background color were too weak. He suggested a white background. The second (88%) did see all colors fine, except he suggested bigger fonts. The hard of hearing person did not have any of the visual problems. All users were confused with some of the controls, which shapes and colors were not consistent. Moreover, they suggested that all management controls should be on the first tab sheet of the application (Fig. 2 – a).

Our unstructured interview revealed a rich set of information which described the following areas: application (m-house), m-learning and mobile technology and the evaluation process. Regarding m-house, one user pointed out that the help functionality was not provided in the application. Another said that lights buttons could be the same size. The next area disclosed some of the interesting questions, for example "would you use this application in your house". Two users absolutely agreed, if the technology would be stable, the other one was more of traditionally oriented and said no or perhaps in the future. Users see the m-technology as an interesting way of controlling devices remotely via wireless internet. Since, the

evaluation methods were not straining they applaud their use. Because they have been aware of the microphone and audio taping the evaluation was intrusive. However, this lasted only for several seconds.

6 Conclusions

The main potential of Internet services; i.e. their universality and accessibility, can be effectively used also for mobile learning. Accessibility can be stated as one of the key factors for planning and designing different educational applications, especially for people with special needs. Moreover, accessibility represents an area that supports the effective use of interactive mobile applications.

Basic guidelines for designing m-learning applications for persons with special needs were given within this contribution. Special attention was given to effective human-computer interaction by applying different HCI factors. Having in mind the guidelines for designing and implementing the m-learning systems we tested the possibility of using m-learning for persons with disabilities by planning and designing the system named m-house. The elements of a wooden model are controlled by micro-controlled system implemented by LabView program using wireless Internet. Specially designed user interface enables direct remote control of different units integrated into the model of the house. User interface for mobile systems, in this case designed for a PDA, served as a platform for examining the functionalities of a micro-controlled system. Furthermore, two usability testing methods, thinking-aloud method and unstructured interview, involving some persons with special needs, were conducted in order to examine the usability of the system. The results gained are valuable for designing improved m-learning applications in the future.

We can conclude that m-learning can be also used for people with special needs. However, the example presented in this contribution, also shows some limitations of mobile devices, such as small display area, limited working memory and restricted use of multimedia contents, which affect the usability of the system when employed by the people with special needs. These limitations, therefore, should be considered precisely when planning and designing the future-oriented m-learning applications for persons with different kinds and levels of impairments.

Acknowledgments. This work has been performed in the framework of European EQUAL Project "Equal Study" run in the period 2004-2007. National co-financier was the Slovenian Ministry of Labour, Family and Social Affairs. Moreover, we would like to thank student Milan Rotovnik for development of m-house model.

References

1. Schulz-Zander, R., Büchter, A., Dalmer, R.: The role of ICT as a promoter of students' cooperation. Journal of Computer Assisted Learning 4, 438 (2002)
2. HSDPA. Source: http://www.umtsworld.com/technology/hsdpa.htm
3. Kulik, C.C., Kulik, J.A., Shwalb, B.J.: The effectiveness of computer applications: a meta-analysis. Journal of Research on Computing in Education 27, 48–61 (1994)

4. Chou, S.W., Liu, C.: Learning effectiveness in a Web-based virtual learning environment. a learner control perspective Journal of Computer Assisted Learning 1, 65 (2005)
5. Kelly, B. et al.: Forcing standardization or accommodating diversity?: a framework for applying the WCAG in the real world. In: Proceedings of the 2005 International Cross-Disciplinary Workshop on Web Accessibility (W4A), Chiba, Japan pp. 46–54 (2005)
6. W3C (World Wide Web Consortium) Source: http://www.w3.org
7. Koyani, S.J., Bailey, R.W., Nall, J.R.: Research-Based Web Design and Usability Guidelines. Source: http://usability.gov/pdfs/guidelines.html
8. Clark, J.: Building Accessible Web Sites. New Riders, Indianapolis (2002)
9. Coyne, K., Nielsen, J.: Beyond ALT text: making the web easy to use for users with disabilities. Nielsen Norman Group, Fremont, CA (2001)
10. Badre, A.N.: Shaping Web Usability: Interaction Design in Context. Addison Wesley Professional, Boston, MA (2002)
11. ATRC - Adaptive Technology Resource Center, A-Prompt: Web Accessibility Verifier, University of Toronto (2006) Source http://www.aprompt.ca/
12. CAST - Center for Applied Special Technology Bobby v5.0. Source: http://www.cast.org/
13. Truwex – Truwex Online check for web accessibility, privacy, quality. Source: http://checkwebsite.erigami.com/accessibility.html
14. SEC508- United States Government, Rehabilitation Act of 1973 (amended in 1998) Section 508. Source: http://www.section508.gov
15. WAI - World Wide Web Consortium (W3C) Web Accessibility Initiative (WAI). Source: http://www.w3.org/WAI/
16. ISO - International Organization for Standardization http://www.iso.org/
17. ITU – International Telecommunication Union. Source: http://www.itu.int/
18. WAI1 - Web Accessibility Evaluation Tools. Source: http://www.w3.org/WAI/
19. Marcus, A.: Principles of effective visual communication for graphical user interface design. In: Baecker, R.M., Grudin, J., Buxton, W.A., Greenberg, S. (eds.) Human-computer interaction: toward the year 2000, 2nd edn. Morgan Kaufmann, San Francisco (1995)
20. LabView. Source: http://www.ni.com/labview/
21. Nielsen, J.: Usability Engineering. Morgan Kaufmann, San Francisco (1994)

Accessible and Adaptive e-Learning Materials: Considerations for Design and Development

Matjaž Debevc, Zoran Stjepanovič, Petra Povalej, Mateja Verlič, and Peter Kokol

University of Maribor, Smetanova 17, 2000 Maribor, Slovenia
{matjaz.debevc,stjepanovic,petra.povalej,mateja.verlic,
kokol}@uni-mb.si

Abstract. The aspect of accessibility and adaptivity is important for future of e-Learning applications. Creating e-Learning applications for everybody, including people with special needs, remains the question. The problem with development of e-Learning applications for everybody is that learner ability and weaknesses are usually neglected as important factors while developing applications. Most of nowadays applications offer lots of unclear information, unsuitable contents and non-adapted mechanisms. This paper suggests basic guidelines for successful design and structuring accessible and adaptive e-Learning applications that consider the requests and needs of people with special needs. It provides an example of design and realization of e-Learning application for receiving ECDL certificate, which includes easy adaptivity and basic accessibility factors. Experimental results of usability testing and pedagogical effectiveness have shown that material, designed following these guidelines, is appropriate and that there must be extra attention paid to learnability factor in the future.

Keywords: e-Learning, accessibility, adaptive user interfaces, usability, evaluation, people with special needs.

1 Introduction

With the development of computers and related technologies, information and computing technology (ICT) has broadened its aspects to the educational area. ICT offers support to students while learning, acquiring knowledge and communicating on several levels. To the teachers, ICT offers support in providing necessary material for distance learning, for informing students, communicating with them and evaluating their results.

Researches have shown that learning with ICT is easier and more effective, especially if ICT is used in the educational process. The researches by Parlangeli et al. [1], Kulik [2] and Bosco [3], who compared traditional learning methods with learning by multimedia systems, showed higher effectiveness of learning, if ICT was used as learning framework. Kulik in his research has shown that students spent 70% less time for learning in some cases then by using classical learning methods. This has proven to be valid especially for material, which included animations and video presentations. Chou and Lou [4] and Schulz et al. [5] have also shown in their

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 549–558, 2007.
© Springer-Verlag Berlin Heidelberg 2007

research, that distance learning using web-based virtual learning environment is effective and useful.

Due to this effectiveness of education using ICT, the question arises whether nowadays e-Learning applications are suitable and successful also for education of all the people, including people with special needs. In other words, how should be the outlook and usability of e-Learning applications in the future, in order to be usable for all the people? The problem in development e-Learning applications is the variety of disturbances and impediments, which people with special needs are having, requesting more or less adapted execution of breeding and education with additional help, which can be either technical or of inter-personal nature [6].

In order to control the impact of e-Learning applications onto success of education, we have used modified (adapted) Course Management Systems (CMS) Moodle for European Computer Driving License (ECDL) of e-materials [7] for qualifying poeple with special needs. We have used additional functionality, which shows characteristics of e-Learning applications in the future. This paper presents following items, which are result of experiences and research during this project:

- Basic guidelines for designing and structuring accessible and adaptive e-Learning applications according to the analysis of needs and requirements of the users.
- Possibilities for developing advanced adaptive e-Learning materials, taking into account different adaptivity criteria
- Explanation of the chosen design of e-materials and working principles of the system. The course entitled “Basic Concepts of Information Technologies” as a part of the European Computer Driving License (ECDL) will be presented as an illustrative example.
- Presentation of evaluation methods focused on the object of evaluation, such as: evaluation of usability of e-materials and evaluation of the pedagogical efficiency.

2 Related Works and Project Related e-Learning Framework

The basics of problem solution for improving educational level of people with special needs using e-Learning applications can be found in other projects, like SMILE [8], ShowVideo [9] and Barrierefrei [10]. The results of other research have already shown that computer aided learning for hearing disabled people is effective enough as additional method for training and development of speech capabilities [11]. Following this, e-Learning applications could be additional help while learning when repeating and refreshing, what has been learnt, at home. This makes the student capable to learn and capture information from study material in verbal and multimedia form, which is the closest to him.

However, all projects so far have focused on a certain group of people with special needs and limited functionality. The results of SMILE project can be, for example, used only as a dictionary for hearing impaired people. The Barrierefrei project is the only one, which tries to equally offer material to different people with special needs. The project is suitable for seeing impaired people, but it does not have sign language video included to be suitable also for hearing disabled people.

This has led us to use curriculum that had been defined by the ECDL Foundation as development and testing basis for e-Learning materials in our project. ECDL/ICDL are the global standards for end-user computer skills, offering candidates an internationally recognized certificate that is globally supported by governments, computer societies, international organizations and commercial corporations.

Education process took place in the form of courses, following the method of blended learning [12]. In the year 2006, in total 20.000 unemployed people in Slovenia were educated for computer literacy within the frame of PHARE 2003 Programme; out of that 300 people in our project– 60 of them were people with special needs.

3 Basic Guidelines for Designing and Structuring Future Accessible and Adaptive Learning Management System

For the analysis of needs and requirements of people and for setting the directives for development of e-Learning applications for everybody, including people with special needs we have used analysis methodology for ICT projects defined by European committee [13]. Methodology includes personal interviews with end-users, questionnaires and organization of discussion panels with brainstorming and feedback. For future user interfaces as well as for the materials, there has been general request for e-Learning applications to contain both accessibility and adaptivity, where there must be a strict distinction between these two terms.

The term "accessible e-Learning materials" is used for e-Learning material that is accessible taking into account physical, sensorial and mental capabilities of users. Furthermore, it presumes that the content is presented in a clear and simple way, as well as that simple, understandable navigation between the pages is used.

The term "adaptive e-Learning materials" means that the content of e-Learning material can be adaptively changed according to user's capabilities and activities. Here, the content could be adapted regarding the pre-knowledge, learning curve, history of activities with e-Learning materials, number of errors in questionnaires and on user profile (visual, hearing centered, age, type of disability).

The analysis had given us some basic guidelines for creating adaptive and accessible e-Learning applications.

3.1 Accessible e-Learning Materials - Basic Guidelines

The highest power of World Wide Web is its wideness and universality. Therefore is the accessibility factor one of the most important aspects. Accessibility is the area, which helps ensuring that all the users can use interactive e-content. Planning and design of accessible e-content needs to become integral part of e-content development and must follow the instructions for design, required by people with special needs.

Table 1 presents some of the most important guidelines for design of accessible e-content, which we have determined using analysis of needs and requirements of people with special needs and considering the instructions by Web Content Accessibility Guidelines (WCAG) from Web Accessibility Initiative (WAI) [14].

Table 1. Impairment and technical solutions

Impairment	Solution for e-Learning content development
Partial sight	Higher number of larger images, font scaling option, color change option, additional information for images, increased involvement of audio information. Hardware based accessibility (screen magnifiers)
Blindness	Verbal image and video descriptions, verbal subtitles, verbal navigation, audio text description, hardware based accessibility (screen readers, Braille keyboards)
Hard of hearing	Video subtitles, image and video oriented content, textual communication option
Deaf	Video subtitles, sign language video, additional glossary in sign language, easy reading content
Impaired dexterity (such as arthritis, tremor, lack of limbs)	Hardware based accessibility (keyboards, touch panels)
Impaired cognition (such as memory, concentration)	Logical and easy navigation, very easy reading content, larger number of image material

When designing e-Learning applications for people with special needs, standards for development of accessible web systems, like ISO 16071 guidelines (Guidance for Software Accessibility) and ISO 9241-14 (Ergonomic requirements) [15], need to be taken into consideration in addition to the guidelines presented in Table 1.

Next aspect of development is the trial of realization of one system for all different users. This case requires system adaptivity in future e-Learning applications.

3.2 Adaptive e-Learning Material - Basic Guidelines

Adaptive e-Learning materials in our project are defined as online learning system, which automatically adapts to students or other users in intelligent way. The techniques mostly originate from artificial intelligence theory and are used for the needs of teaching learning systems to adapt individual needs and user wishes. According to this, adaptive systems can further be split into subsections, depending on who gives initiative, proposal, decision and execution of activities for the change [16]. Analysis of users' requests have shown, that full adaptivity is not suitable for changing the contents of learning material, since it may cause losing a connection between the user and the content and structure, if the content gets automatically changed without user's consent.

Adaptive web based educational systems usually include expert modules, student modules, tutoring modules and user interface module, all designed based on different criteria, such as: pre-knowledge, learning curve, history of activities with e-Learning materials, and item response theory [17]. As result of connecting certain (students and learning) and these modules, adaptivity of e-Learning application is offered to users.

The question, what can be changed in the e-Learning application, remains. According to Brusilovsky [18], the adaptive presentation and adaptive navigation can be done in following ways in hypermedia supported learning materials:

- *Adaptive presentation* presents option for the content of hypermedia designed web page to adapt in sense of user's goals, knowledge and other information that is stored about the user (student's model). In general, it is required that less experienced users get offered more simple content, whilst advanced users get presented with additional, in depth information.
- The goal of *Adaptive navigation* is to support student in orientation and navigation using visual hyperlinks. Adaptive learning system sorts, emphasizes or partially hides links to the web page, so that the user can easier access it. This approach can be implemented in very effective and natural way.

The basic task of development of adaptive e-Learning applications is to create simple, efficient and effective learning material. However, special attention needs to be paid into creation of adaptive systems, so that the user gets offered only adaptivity of one small part of whole course management system (e.g. navigation or contents), otherwise the user may lose the clarity and navigability of learning process.

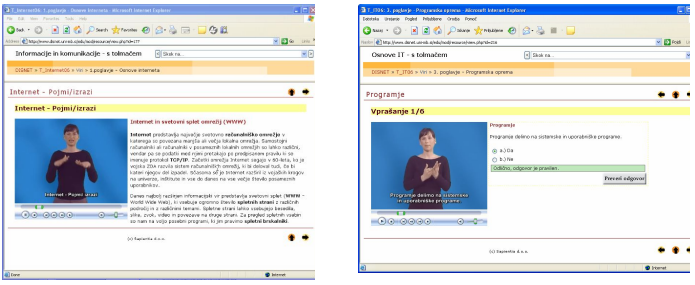
4 Accessible and Adaptive e-Learning Application - ECDL e-Learning Example

Having in mind the guidelines for designing and implementing e-Learning systems for people with special needs listed in previous section we decided to present the content in the form of web (HTML) pages inside of Moodle CMS system [19]. We intentionally kept the design simple and avoid the complex graphic elements in order to reduce the influence of disturbing factors that could distract attention from the content.

We designed basic adaptivity, in which system offers initiative and proposal; other activities (decision and execution) were carried out by the user. System therefore offers the content within individual topics, which is divided according to two defined levels: basic and advanced. The basic levels were designed for users having only basic computer skills or even complete beginners. The advanced level was designed for users, experienced in basic work with the computer with a goal to upgrade their knowledge with new information. Users who gained/repeat the knowledge foreseen for the basic level quicker could also see the chapters at the advanced level and prepare themselves for final ECDL exam.

According to accessibility guidelines, we have designed the materials in such a way that particular web page was divided into two screens (Figure 1a). Left hand side shows sign language video and the right hand side simple designed content with basic navigation buttons.

Next aspect of accessibility was implemented with the use of simple and usable navigation inside of e-Learning material. In this way, the user always knows in which module one works and which activities are being used. Important aspect at the accessibility design was the loading speed of web pages, especially video. Therefore video includes marker tags and was segmented into smaller sections that correspond



a) b)

Fig. 1. ECDL for deaf and hard of hearing people – a) classical view of material with sign language interpreter b) material with question for individual knowledge testing

to the units the material was divided into. This solution has extremely fastened video and web page loading time. Assessment of the knowledge can be stated as a key factor for successful and effective learning. Therefore, at the end of each lesson we include short questions (Figure 1b). The answers were not registered and the user could answer the questions several times. After answering, the system responded whether the answer was correct or not. If reasonable, there was also a short explanation in case of an incorrect answer.

5 Pedagogical and Usability Testing - Results

The decision of suitable evaluation is the key point for testing educational system. Since we were dealing with online classes and based on investigation on other researches, like Squires and Preece [20], Ardito et ell [21], Dringus and Cohen [22], Holzinger [23] and Achtemeier [24], we have used evaluation methods mostly depending on the object of evaluation, like evaluation of user-friendliness of the system, general impression, pedagogical effectiveness and effectiveness of each class.

5.1 SUMI Evaluation

The usability of the system was tested by 10 hard of hearing trainees. As per instructions in Software Usability Measurement Inventory (SUMI) [25] software documentation, a sample of 10 or more users per system being evaluated is required to get stable statistical results. Four participants were males and six of the participants were females. All participants had no basic computer and web browser experience before joining the e-course. The results of SUMI evaluation are shown in Figure 2. Data are presented in terms of the mean, the 95% upper and lower confidence levels (UCL and LCL). These descriptive statistics are given for the global usability scale and each of the five usability sub-scales (Efficiency, Affect, Helpfulness, Control, Learnability).

The global scale, the most reliable of all the SUMI scales, indicated that the usability of our user interface is comparable to successful commercial systems. In terms of the usability, sub-scales show that results are in general consistent, and nearly all fall within the desired range. The exception is a learnability sub-scale which

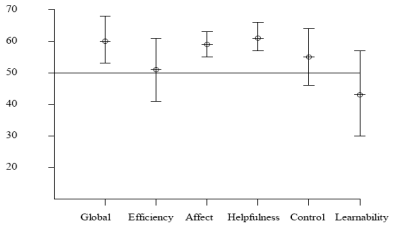


Fig. 2. Comparison of global and subscale usability scores (SUMI – evaluation)

falls under the desired range. Participants did not share the same opinion about application's efficiency, and had some problems with control of the application. Lower efficiency level shows that users may not always know what to do next or the software may sometimes work in a strange, inconsistent way. High helpfulness level on the other side shows that this application communicates clearly, users can understand the way it works and it gives them helpful hints and instructions.

The outcome of SUMI questionnaire confirms our statement, that the system is usable for integration into educational process. Detailed analysis of answers to questions leads to conclusion, that once the users overcame the initial difficulties with the system, they found the experience of working with it a relatively pleasant, with helpful support and reasonable performance.

5.2 Pedagogical Effectiveness

In the framework of total evaluation of the project, our aim was to evaluate pedagogical effectiveness of the classes. Chosen methodology is based on Sonwalkar method [26]. Sonwalker talks about Pedagogy Effectiveness Index (PEI), which consists of 3 dimensions (Media Types, Learning Styles in Interactions) that are further composed of several factors and this results in suitable probability distribution matrix.

Since our aim was to evaluate pedagogical effectiveness of classes for adults with special needs, we needed to adapt Sonwalker's probability matrix and the procedure. The result was new methodology for evaluation of PEI, which considers additional factors for people with special needs. We have added new aspect to learning styles, namely perception styles (visual, kinesthetic and audial), which matches different styles of impairment. In media styles, organization of material, in accordance with special needs, was added. Finally, a tutor was introduced into interaction, to receive feedback from students about the qualification of mentor to work with people with special needs.

The presence and good representation of each factor was evaluated by participants. Based on this evaluation, the PEI was calculated and it provided us with the information, whether the classes were pedagogically successful or not for the participants. 116 participants were involved, who were split into 5 age groups. 81 participants were females, 34 were males. Pedagogical index was calculated based on the equation, which returns a value between 0 and 1, whereas 1 is the best pedagogical effectiveness. The final result for all the participants was 0.68, meaning that participants were generally satisfied with the pedagogical work.

5.3 Course Effectiveness

In order to evaluate the efficiency of individual course, we implement pre-exam before the course where we tested the trainee for ones prior knowledge and post-exam, where the trainee took a final test after he finished the course. Each trainee had one week for learning each module. The pre-exam results were compared to post-exam results in order to establish how much did the trainees learn in one week. The results showed that trainees demonstrated very good response to using e-Learning materials. On the average they improved their knowledge for 25,5 % through all four modules.

The best results were obtained in practical modules (Word processing and Windows) where the average result on final exam was almost 80%. The trainees experienced some difficulties in theoretical module since the interpretation of many basic notions in sign language were very difficult. Therefore the average result was almost 70%. In all modules the gain of knowledge was significant.

6 Conclusions

Information and computing technology offers today the students effective support while learning, acquiring knowledge and communicating on several levels. Using ICT, teachers can provide necessary material for e-learning; furthermore they can effectively communicate with their students. Due to numerous advantages, e-Learning applications are suitable and successful also for education of people with special needs. However, the problem related to the development of such e-Learning applications is the variety of disturbances and impediments, which people with special needs are having, requesting more or less adapted execution of breeding and education with additional help, which can be either technical or of inter-personal nature.

In our work we have used adapted Course Management Systems Moodle for acquiring European Computer Driving License using e-materials for people with special needs In order to control the impact of e-Learning applications onto success of education. Additional functionalities, which show characteristics of e-Learning applications in the future, have been developed. For the analysis of needs and requirements and for setting the directives for development of e-Learning applications for everybody, including people with special needs, we have used analysis methodology for ICT projects defined by the European committee. The methodology included personal interviews with end-users, questionnaires and organization of discussion panels with brainstorming and feedback. For future user interfaces as well as for the materials, there has been general request for e-Learning applications to contain both accessibility and adaptivity, where there must be a strict distinction between these two terms. First of all, e-Learning material must be accessible taking into account physical, sensorial and mental capabilities of users. Furthermore, it presumes that the content is presented in a clear and simple way, and that simple, understandable navigation between the pages is used. In future-oriented adaptive e-Learning materials the e-Learning content can be adaptively changed according to user's capabilities and activities taking into account the pre-knowledge, learning curve, previous experiences with e-Learning materials, number of errors in

questionnaires and user profile, resp. type of disability. Our research work resulted with analysis that had given us some basic guidelines for creating effective adaptive and accessible e-Learning applications.

Conclusion can be made that future-oriented adaptive web based educational systems will include expert modules, student modules, tutoring modules and user interface module, all designed based on different criteria. In our work we designed basic adaptivity, in which system offers initiative and proposal; other activities were carried out by the user. For pedagogical and usability testing we have used evaluation methods mostly depending on the object of evaluation, such as evaluation of user-friendliness of the system, general impression and pedagogical effectiveness. SUMI evaluation was used for assessing the usability of the system. The outcome of SUMI questionnaire confirms that the system is usable for integration into educational process. We also evaluated the pedagogical and course effectiveness using appropriate methods; both were valued above the average. We can conclude that when planning and developing the next generation of future-oriented applications, accessibility and adaptivity of e-Learning materials must be considered as critical success factors.

Acknowledgments

This work has been performed in the framework of European PHARE 2003 project DISNET entitled: “Improvement of computer literacy of unemployed adults” run in the year 2006. Fundamental to the study was the close collaboration with Maribor Adult Education Centre (AEC Maribor).

References

1. Parlangeli, O., Marchigiani, E., Bagnara, S.: Multimedia systems in distance education: effects of usability on learning. Interacting with Computers 12, 37–49 (1999)
2. Kulik, C.C., Kulik, J.A., Shwalb, B.J.: The effectiveness of computer applications: a meta-analysis. Journal of Research on Computing in Education 27, 48–61 (1994)
3. Bosco, J.: An analysis of evaluations of interactive video. Educat. Techn. 25, 7–16 (1986)
4. Chou, S.W., Liu, C.: Learning effectiveness in a Web-based virtual learning environment: a learner control perspective. Journal of Computer Assisted Learning 21, 65–76 (2005)
5. Schulz-Zander, R., Büchter, A., Dalmer, R.: The role of ICT as a promoter of students’ cooperation. Journal of Computer Assisted Learning 18, 438–448 (2002)
6. Smith, C., Mayes, T.: Telematics Applications for Education and Training: Usability Guide. ICBL (1996)
7. Carpenter, D., Dolan, D., Leahy, D., Sher-Wood-Smith, M.: ECDL/ICDL: a global computer literacy initiative. In: the Proceedings of Conference on Educational Uses of Information and Communication Technologies (ICEUT), Beijing: Publishing House of Electronics Industry (PHEI) (2000)
8. Kronreif, G., et al.: SMILE: demonstration of a cognitively oriented solution to the improvement of written language competence of deaf people. In: the Proceedings of the 7th International Conference on Computers Helping People with Special Needs-ICCHP’2000. Karlsruhe, Germany (2000)

9. Vanderheiden, G.C.: Full Visual Annotation of Auditorially Presented Information for Users Who Are Deaf: ShowSounds. In: the Proceedings of the RESNA International Conference, Toronto, Canada (1992)
10. ECDL-Barrierefrei. Source: http://www.barrierefrei.ecdl.at
11. Druin, A., Hendler, J.: Robots for Kids: Exploring new technologies for learning. Morgan Kaufmann, San Francisco (2000)
12. Osguthorpe, R.T., Graham, C.R.: Blended Learning Environments: Definitions and Directions. Quarterly Review of Distance Education 3, 227–233 (2003)
13. British Educational Communications and Technology Agency (Becta): Website accessibility guide – Understanding and implementing accessibility (2005)
14. Kelly, B., Sloan, D., Phipps, L., Petrie, H., Hamilton, F.: Forcing standardization or accommodating diversity: a framework for applying the WCAG in the real world. In: Proceedings of the 2005 International Cross-Disciplinary Workshop on Web Accessibility (W4A), Chiba, Japan pp.46–54 (2005)
15. Gulliksen, J., Harker, S., Vanderheiden, G.: Guidelines, standards, methods and processes for software accessibility. International Journal Universal Access in the Information Society 3, 1–5 (2004)
16. Schneider-Hufschmidt, M., Kühme, T., Malinowski, U. (eds.): Adaptive User Interfaces: Principles and Practise. Elsevier, North Holland, Amsterdam (1993)
17. Chen, C.-M., Lee, H.-M., Chen, Y.-H.: Personalized e-Learning system using Item Response Theory. Computers & Education 44, 237–255 (2005)
18. Brusilovsky, P.: Adaptive Hypermedia. User. Modeling and User.-Adapted Interaction 11, 87–110 (2001)
19. Dougiamas, M., Taylor, P.: Moodle: Using Learning Communities to Create an Open Source Course Management System. In: Kommers P., Richards, G. (eds.) Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications. Chesapeake VA: AACE, pp. 171–178 (2003)
20. Squires, D., Preece, J.: Predicting quality in educational software: Evaluating for learning, usability and the synergy between them. Interacting with Computers 11, 467–483 (1999)
21. Ardito, C., Costabile, M.F., De Marsico, M., Lanzilotti, R., Levialdi, S., Roselli, T., Rossano, V., Tersigni, M.: An Approach to Usability Evaluation of e-Learning Applications. International Journal Universal Access in the Information Society, Special Issue on User-Centred Interaction Paradigms for Universal Access in the Information Society, pp. 1–14 (2005)
22. Dringus, L.P., Cohen, M.S.: An adaptable Usability Heuristic Checklist for Online Courses. In: the Proceedings of the ASEE/IEEE Frontiers in Education Conference (2005)
23. Holzinger, A.: Usability Engineering Methods for Software Developers. Communication of the ACM 48, 71–74 (2005)
24. Achetemeier, S., Morris, L., Finnegan, C.: Considerations for developing evaluations for online courses. Journal of Asynchronous Learning Networks, vol. 7 (2003)
25. Kirakowski, J., Corbett, M.: SUMI: The Software Usability Measurement Inventory. British Journal of Educational Technology 24, 210–212 (1993)
26. Sonwalkar, N.: A New Methodology for Evaluation: The Pedagogical Rating of Online Courses. Syllabus 15, 18–21 (2002)

Web 2.0 Technology: Future Interfaces for Technology Enhanced Learning?

Martin Ebner[1], Andreas Holzinger[2], and Hermann Maurer[3]

[1] Graz University of Technology, A-8010 Graz, Austria
WG Social Learning
martin.ebner@tugraz.at
[2] Medical University Graz, A-8036 Graz, Austria
Institute for Medical Informatics, Statistics & Documentation (IMI)
Research Unit HCI4MED
andreas.holzinger@meduni-graz.at
[3] Graz University of Technology, A-8010 Graz, Austria
Institute for Information Systems and Computer Media (IICM)
hmaurer@iicm.edu

Abstract. Web 2.0 is an emerging catch phrase and the applications associated with it shocked the traditional eLearning world. However, is this really all new and can it be considered suitable for future interfaces for technology enhanced learning? In this paper, we present some experiences with *LearnLand.* This is an application designed and developed at Graz University of Technology and which has been running there since October 2006. It is based on the open-source software ELGG, which is a social software offering a high degree of choice, flexibility and openness and is considered as a system that places people at the hub of the activities. Our experimental research demonstrated that exactly this *ease of use* aspect is an absolute necessity for successful Web 2.0 learning applications. The tools of Web 2.0 have crossed Moore's chasm and reached the early majority, where they are evolving rapidly. However, the idea of *social software* itself, especially in learning scenarios, is not as far developed as we may imagine. It is necessary to take into account the fact that too few innovators and early adopters are actually using Web 2.0 technology to enhance existing learning behaviors. Insufficient educational concepts with Web 2.0 technologies require much future work.

Keywords: Technology Enhanced Learning, Web 2.0, e-Learning.

"Web 2.0 is an attitude – not a technology; this means there is no technological revolution, it is a social revolution." Stephen Downes (2006), [1]

1 Introduction

The buzzword Web 2.0 and all its applications shocked the traditional eLearning World [2]. Blogs, Wikis, Podcasts, RSS, etc. seem to be the future of learning [3].

During this evolution also a new name emerged: *E-learning 2.0* [1]. However, all those technologies are *not new.* Old technology shall lead to an impact in teaching and

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 559–568, 2007.
© Springer-Verlag Berlin Heidelberg 2007

learning behavior? What is the reason for such a revolution? A lot of questions appear. Especially in higher education experiences in using such new technologies are lacking and the often propagated *social learning* happens rarely; taking into consideration that social learning basically results from a number of social phenomena, the most important of which is *social facilitation* [4].

According to Berners-Lee (1989) [5], three major components changed the use of the Internet:

Accessibility for all: Today broadband access to the internet is provided area wide in all Western countries and it is *affordable* for nearly everybody. People can use it for any kind of information collecting, shopping, selling etc.
Devices: Similarly, technological devices are both available and affordable and it is possible to connect with the internet with these mobile and pervasive devices. Although, Personal Computer (PC) is still the main access point to the Internet, WLAN, GPRS etc. are considerably increasing. Pervasive computing is an emerging area, where technology more and more disappears, as Mark Weiser suggested already in 1991 [6].
Usability: Concerning the issues above, the *ease of use* factor is definitely the most important one [7]. In the past, only expert knowledge guaranteed the contribution to the Web. A typical end user without any HTML skills was unable to contribute.

It is interesting to note, that within education the use of computers has mostly been focused on enhancing learning in formal settings, typically in the traditional classroom or computer lab [8]. However, learning does not always take place within such formal learning settings and much learning does not happen in the traditional way, which means that e.g. incidental learning, which is often ignored, is of vital importance [9]. The use of mobile devices could even expand learning possibilities and solve the problem of being tied to a particular location [10]. Generally, the combination of e-learning and mobile computing is called mobile learning (m-learning) and promises the access to applications that support learning anywhere, anytime. Today, sophisticated and functioning hardware is available, whereas innovative, affordable, adaptable and most of all *usable interfaces* remains still the greatest challenge [10], [11].

In 2004 O'Reilly organized the first Web 2.0 Conference and the following slogans emerged: "The user is the content", "Internet is about people", "Social revolution", "Social Networking" etc.; Flickr, YouTube or MySpace were presented as *the* platforms of the next generation Web. It is amazing how fast the use of the Web changed and it is remarkable that all involved research fields, including education, are currently struggling with the incredible speed of technological development.

2 Theoretical Background

Discoveries about learning, including learning conditions and learning success, can be systematically and analytically combined into so called *Theories of Learning.* These contain the basic concepts, according to which educational software functions. It is interesting that mostly these theories are not taken into consideration, due to the fact that many technologists are unaware, or even not interested in those. However, at the same time, educational experts, psychologists and pedagogists, are often disinterested

in computer science and/or lack the necessary basic technological knowledge, hence interdisciplinary thinking is of greatest importance in this field [12].

Before discussing, what Web 2.0 probably can change, it must be pointed out that learning will not be easier as before, because learning is an active cognitive process on the part of the learner, where knowledge and understanding is constructed by the learner [13], [14], [12]. However, learning is also a *social process* and proceeds by and through conversation [15] and interaction [16]. Exactly these issues could be a big advantage of Web 2.0 technologies in the future. A large number of didactical scenarios and technical tools have been developed [17]. However, many projects showed lacks concerning successful interaction between the learners and an increasing learning outcome [16]. This implicates that the application of learning theories (behavioristic, cognitive and constructivistic) is necessary in order to strengthen the educational design.

There is some evidence that learners perform better in teams [18], [19], and if learners are part of the educational process and have the possibility to interact and to communicate [20]. However, it is necessary that facilitators engage learners into meaningful interaction and to overcome the transactional distance of online learning [21].

According to Moore, it is necessary to 1) interact with the *content,* 2) interact with the *instructor* and to interact with *other learners* [22]. However, it is necessary to include a fourth type of interaction: to interact with *(new) technologies* [23]. Consequently, good Usability is a central concern of these new technologies, in order to be helpful for the learners.

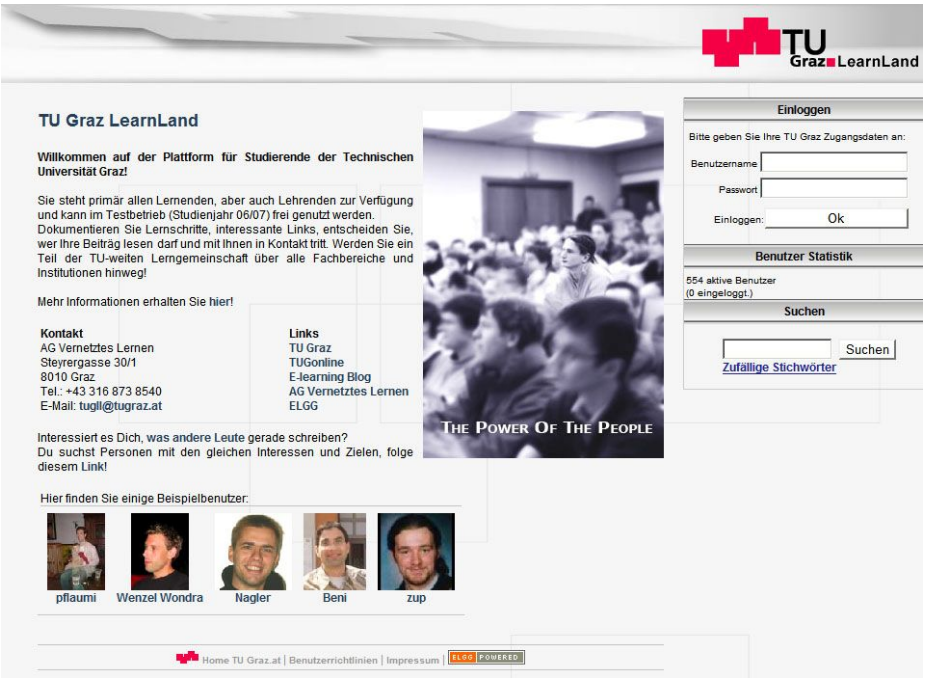


Fig. 1. Screenshot of LearnLand

3 The Project LearnLand

LearnLand is in operation since October 2006 at Graz University of Technology and is accessible to every students and teachers university wide (a screenshot of the start page can be seen in figure 1). LearnLand is based on the open source software ELGG (http://elgg.org – last visited: 22.12.2006) and is described by the developers as *"open source software platform designed to allow people to easily connect and share resources"*. ELGG is comparable with a blogosphere for example as blogger.com, however, it provides more possibilities with have a strong focus on learning and learning in communities [24].

4 LearnLand Technical Background

Both ELGG and LearnLand are based on a modular PHP structure, however, the architecture is not object orientated, due to the fact that modifications can also be done by people who are not experts in PHP programming. The main benefits of this system are the *ease of use* and *easy extensibility* by the administrators. Technically, XML files (RSS, FOAF, Atom) – which form a fundamental basis of Web 2.0 – are funneled through the template-engine; this enables an easy modification of the design.

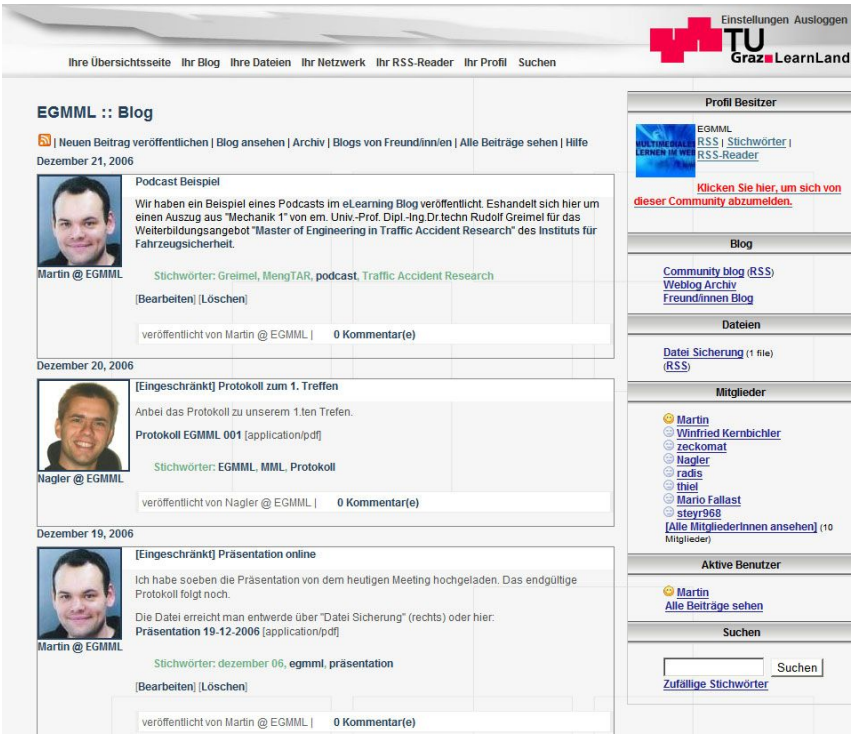


Fig. 2. Example of LearnLand: Community Weblog

A mySQL database is used for storage of all contributions and the user specific data. The hardware requirements are low. An advanced feature of LearnLand is the synchronization of username and password with the central administration system of the University. Security is guaranteed by using LDAP connections and SSH protocols.

The following components are part of the standard installation of LearnLand:

Weblog. Every user gets his or her individual and personal Weblog (see figure 2). End users can provide their personal content by assistance of a simple WYSIWYG-Editor. The view of each contribution can be restricted, in order to make accessible the content only to selected end users; only selected end users can read and comment the content. The use of keywords allows easy search, which is also part of the so called tagged clouds. The use of HTML is possible and the use of hyperlinks or data files from the data pool is in the same way possible as embedding external objects with the object-tag. All things which can be done with usual homepages are possible.

Data pool for the end users. Every end user is provided with the possibility to upload files which they would like to save or share with other end users. Implementing user rights or to admit keywords is comparable to the Weblog. Each file can be connected with a blog contribution or shared with specific users or communities.

Personal Information. The Blogsphere can also be used to provide a short presentation of the end user. This is the main idea of ePortfolios ; in combination with blog contribution this is a core functionality of the future.

Community building. Each user can create his or her individual community. For this community an individual Weblog is created, in order to present the results of teamwork and to share.

Keyword Tagging. One of the most important features is the implemented keyword tagging. This means if a keyword is used once more, independently from the end user, the software automatically links all matching contributions. Consequently, end users with similar interests or similar problems can match easily (community building).

RSS-Reader. Every Weblog gets its individual RSS-Feed and each RSS-Feed from any other webpage can be displayed within. Additionally, it is possible to embed a RSS-Feed in a LearnLand Blog directly. Consequently, external Weblogs (for example a typical Wordpress Blog) can be shown within the environment; the blogger have to blog only once – in their private blog – therefore the contribution is shown in multiple ways (both in their private blog and in the learning environment).

Personal Information. Every participant is able to edit and publish any kind of private data. Of course, they can check if and when some entries are provided and accordingly they can change and adapt the visibility to other end users.

5 Usability Tests

Experimental Setting. Keeping in mind, that one of the crucial factors of Web 2.0 applications is the *ease of use*, we performed a series of usability tests. To gain first experiences with typical end users, we applied standard inspection methods including thinking aloud tests and video analysis of end user behavior [25]. Together with the

programmers and administrators we performed three evaluation rounds with five different end users each (N=15). After every evaluation round the design was adapted according to the findings and the new version was evaluated within the next evaluation round.

5.1 Results

First round: Ten tasks were presented to each of the five end users. Typical tasks included log on, produce a blog contribution, search for a person etc.; the given time to accomplish these tasks was 30 minutes. All tasks were part of the core functionality, however, on average, only 4 out of 10 tasks (40%) were fulfilled.

Second round: The second group got exactly the same tasks as the first group. The only difference was that meanwhile the interface was adapted, according to the lessons learned within the first round, consequently to make the tasks more usable. On average 9 out of 10 tasks (90%) were fulfilled.

Third round: Due to the high rate of improvement during the second round, we decided to provide the end users within this round with ten *new* tasks; these included advanced functionalities, for example file sharing with particular other end users. The gained results provided insights into end user behavior and resulted in revised version of the interface.

6 Discussion

The rapid usability tests showed that the *ease of use* aspect is absolutely necessary for successful use of a Web2.0 application interface. Immediate fixing the problems and rapid reiteration of the interface improved the interface of LearnLand significantly.

Based on the lessons learned with end user behavior during the usability tests, the following additional implementations were made:

Improvement of the community feeling: Online end users are displayed in the sidebar and also their contributions are now displayed immediately after their log in.

Improvement of the search function: One of the disadvantages of the folksonomy, which is, however, also an advantage in our specific case, is the missing hierarchy; end users felt very comfortable that other end users can allocate individual keywords without thinking about (often difficult) categories; if the pool of data increases and in relation to the learning context it is easier to have pre-selected data. Consequently, we implemented categories in order to allow a certain classification.

Improvement of data collection: It was perceived as useful to save hyperlinks of the end users and to share with other end users. Supported by a ranking system, the quality of the bookmarks can be ensured. This feature is comparable to common social bookmarking tools as del.icio.us, but focused on educational requirements.

Concerning the future of learning in combination with Web 2.0 applications, we emphasize that the technological developments are quite advanced, however, the key for designing future interfaces for technology enhanced learning is *ease of use.* This needs the concentration of research on the end users, however, the findings must be

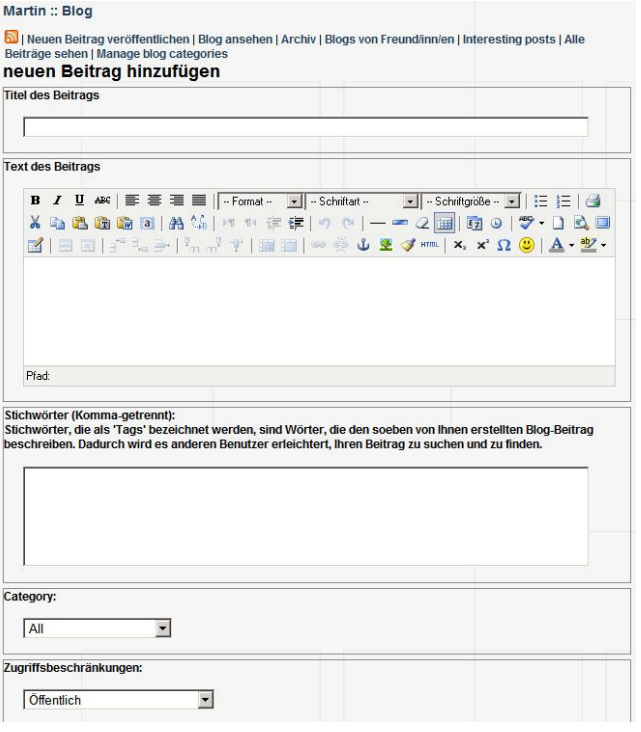


Fig. 3. A view to the Editor of LearnLand

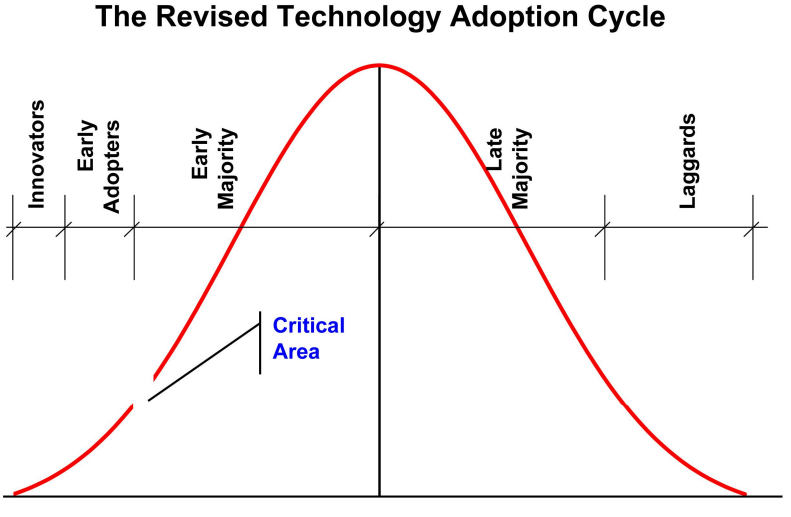


Fig. 4. Technology Adoption Life Cyle, adopted to Web 2.0 technologies according to [27]

integrated at technological level [26]. The challenge of our work is to think about **which** and **how** the combination must be done, in order to address the usability aspects.

Finally, it must be pointed out that the interviews after the usability tests turned out one very interesting point: Whether or not the participants were able to fulfill nearly all tasks in the end of the series, they did not totally understand how they can apply LearnLand in order to support their learning processes, i.e. using Web 2.0 technologies for learning purposes is not as logical as we might believe. This leads us to take into consideration the famous Technology Adoption Life Cycle (compare with figure 4) by Moore [27]. He distinguished between Innovators, Early Adopters, Early Majority, Late Majority and Laggards relating to high-tech products. We expand this view to Web 2.0 technologies, especially to *social software* tools (see figure 4).

Web 2.0 technologies have crossed the chasm or reached the Early Majority. The increasing importance and availability of the applications are evident, however, the idea of social software itself, especially in learning scenarios, is not yet as far as we might believe. Some innovators and early adopters are actually using Web 2.0 technologies to enhance existing learning behaviors, however, insufficient educational concepts with Web 2.0 technologies require much future work.

7 Conclusion and Future Work

We can see two major directions for the future – a technological and a social. Without doubt, developing *future interfaces* will be a necessity; however the typical end user – the learner – must be in the centre of all our work. The *ease of use* aspect, which includes the *accessibility*, is one of the strengths of Web 2.0 technologies. Nowadays, we need to think about **what** we write and not **how** we write, the content is the most important issue. This would be the true revolution in Web 2.0 work. Bearing in mind that only hard usability work leads to user acceptance, Graz University of Technology has strongly focused on this area [28].

A further issue is the contrast between the technological growth of Web 2.0 and the evidently slower increase in tool usage and acceptance. This is the so called social component. One chance of proper use of these technologies is to fight information overload, consequently our future work must concentrate on **integrating** technology with modern and psychological sound didactic principles.

Acknowledgements

We like to express our gratitude to Behnam Tharagi, student of telematics of Graz University of Technology, for supporting this study. We thank also the whole team of Social Learning and the team of the computing department for their endless patience during the installation and implementation process. Finally, we like to express our gratitude to all students participating within this project and providing the necessary content.

References

1. Downes, S.: E-learning 2.0. ACM eLearn Magazine, (October 2005) vol. 10 (2005)
2. O'Reilly, T.: Web 2.0: Stuck on a name or hooked on value? Dr. Dobbs Journal 31(7), 10–10 (2006)
3. Richardson, W.: Blogs, Wikis, Podcasts, and Other Powerful Web Tools for Classrooms. Corwin Press, Thousand Oaks (CA) (2006)
4. Conte, R., Paolucci, M.: Intelligent social learning. Jasss-the Journal of Artificial Societies and Social Simulation 4(1), U61–U82 (2001)
5. Berners-Lee, T.: Information Management: A Proposal. /proposal.html (last access: 2007-01-31) (1989) http://www.w3.org/History/
6. Weiser, M.: The computer for the twenty-first century. Scientific American 265(3), 94–104 (1991)
7. Ziefle, M.: The influence of user expertise and phone complexity on performance, ease of use and learnability of different mobile phones. Behaviour & Information Technology 21(5), 303–311 (2002)
8. Mifsud, L.: Alternative Learning Arenas - Pedagogical Challenges to Mobile Learning Technology. In: EducationIEEE International Workshop on Wireless and Mobile Technologies in Education (WMTE'02), pp. 112–117 (2002)
9. Holzinger, A., Pichler, A., Maurer, H.: Multi Media e-Learning Software TRIANGLE Case-Study: Experimental Results and Lessons Learned. Journal of Universal Science and Technology of Learning 0(0), 61–92 (2006)
10. Holzinger, A., Nischelwitzer, A., Meisenberger, M.: Mobile Phones as a Challenge for m-Learning: Examples for Mobile Interactive Learning Objects (MILOs). In: Tavangarian, D (ed.) 3rd IEEE PerCom, pp. 307–311 (2005)
11. Holzinger, A., Nischelwitzer, A., Meisenberger, M.: Lifelong-Learning Support by M-learning: Example Scenarios. ACM eLearn Magazine, vol. 5 (2005)
12. Holzinger, A.: Multimedia Basics, Volume 2: Learning. Cognitive Fundamentals of multimedial Information Systems. Laxmi, New Delhi (2002) http://www.basiswissen-multimedia.at
13. Wittrock, M.C.: Learning as a generative process. Educational Psychologist 11, 87–95 (1974)
14. Papert, S., Harel, I.: Constructionism. Ablex Publishing, Norwood (NJ) (1991)
15. Motschnig-Pitrik, R., Holzinger, A.: Student-Centered Teaching Meets New Media: Concept and Case Study. IEEE Journal of Educational Technology & Society 5(4), 160–172 (2002)
16. Preece, J., Sharp, H., Rogers, Y.: Interaction Design: Beyond Human-Computer Interaction. Wiley, New York (2002)
17. Collis, B.: New didactics for university instruction: why and how? Computers & Education 31, 373–393 (1998)
18. Egerman, K.: Effects of Team Arrangement on Team Performance - a Learning-Theoretic Analysis. Journal of Personality and Social Psychology 3(5), 541 (1966)
19. Neufeld, D.J., Haggerty, N.: Collaborative team learning in information systems: A pedagogy for developing team skills and high performance. Journal of Computer Information Systems 42(1), 37–43 (2001)
20. Friesen, N., Anderson, T.: Interaction for lifelong learning. British Journal of Educational Technology 35(6), 679–687 (2004)

21. Sargeant, J., Curran, V., Allen, M., Jarvis-Selinger, S., Ho, K.: Facilitating interpersonal interaction and learning online: Linking theory and practice. Journal of Continuing Education in the Health Professions 26(2), 128–136 (2006)
22. Moore, G.E.: Three types of interaction. The American Journal of Distance Education, vol. 3(2) (1989)
23. Bouhnik, D., Marcus, T.: Interaction in distance-learning courses. Journal of the American Society for Information Science and Technology 57(3), 299–305 (2006)
24. Kolbitsch, J., Maurer, H.: The Transformation of the Web: How Emerging Communities Shape the Information. We Consume Journal of Universal Computer Science 12(2), 187–213 (2006)
25. Holzinger, A.: Usability Engineering for Software Developers. Communications of the ACM 48(1), 71–74 (2005)
26. Kolbitsch, J., Maurer, H.: Community Building around Encyclopaedic Knowledge. Journal of Computing and Information Technology 14(3), 175–190 (2006)
27. Moore, G.A.: Crossing the Chasm, Marketing and Selling High-Tech Products to Mainstream Customer (revised edition). HarperCollins Publishers, New York (1999)
28. Ebner, M., Scerbakov, N., Maurer, H.: New Features for eLearning in Higher Education for Civil Engineering. Journal of Universal Science and Technology of Learning 0(0), 93–106 (2006)

Computer Interface for Learning and Using Virtual Instrument

Yongkai Fan[1], Tianze sun[1], Jun Lin[2], Xiaolong Fu[1], and Yangyi Sui[2]

[1] Computer and Infromation Management Center, Tsinghua Uinversity. 100084, Haidian District, Beijing China
{fyk,stz,fxl}@cic.tsinghua.edu.cn
[2] College of Instrument and Electrical Engineering, Jilin Uinversity. 130026, Changchun, Jilin Province, China
{lin_jun,syy}@jlu.edu.cn

Abstract. Computer interfaces representation, design and implementation as the computer software outward window had a large impact on software learning and using, especially for virtual instrument. Several types of virtual instrument developing software are available for the virtual instrument development and parts of them have great influence in the instrument science field. However, only a limited number of inexperienced or previous untrained people are able to well utilize them. Part of the limitation stems from the difficulty in learning how to use them and part of from the demand of software developing expertise background or hardware design abilities. Therefore, user friendliness of virtual instrument software is needed for a great number of people who are without expertise background or hardware design abilities. There are a number of features in the software described in this article would server to meet the need, such as requirement-driven idea from human-computer interface implementation aspect helped step obstacles encountered by users with limited experience. An experimental interface design has been developed to use an advanced object-oriented development environment, thus allowing a great deal of flexibility in implementing changes and adding new features in order to provide a friendliness operation interface to actual users.

Keywords: Human-computer interface, Virtual instrument, Module instrument.

1 Introduction

Widespread application of personal computer and advent of new technologies in computer science over the past two decades has given rise to a new world in present measurement and instruments filed--virtual instrument. The importance of virtual instrument in measurement and instrument filed results in it as a compulsory course appeared in engineering students' curriculum and as a substitute for traditional instrument [2, 4, 7]. Virtual instrument programming environment, tool, and even graphic languages rush to support the process of learning virtual instrument. The graphical programming software package LabVIEW (laboratory virtual instrument engineering workbench, example of one rare panel of LabVIEW is shown as fig. 1)

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 569–577, 2007.
© Springer-Verlag Berlin Heidelberg 2007

has made the concept of the virtual instrument (VI) a practical reality [1, 3, 5, 6]. LabVIEW gradually becomes the substitute name of virtual instrument. However, it is a difficult subject for beginner to learn and program a virtual instrument, especially one without any related software programming experience.

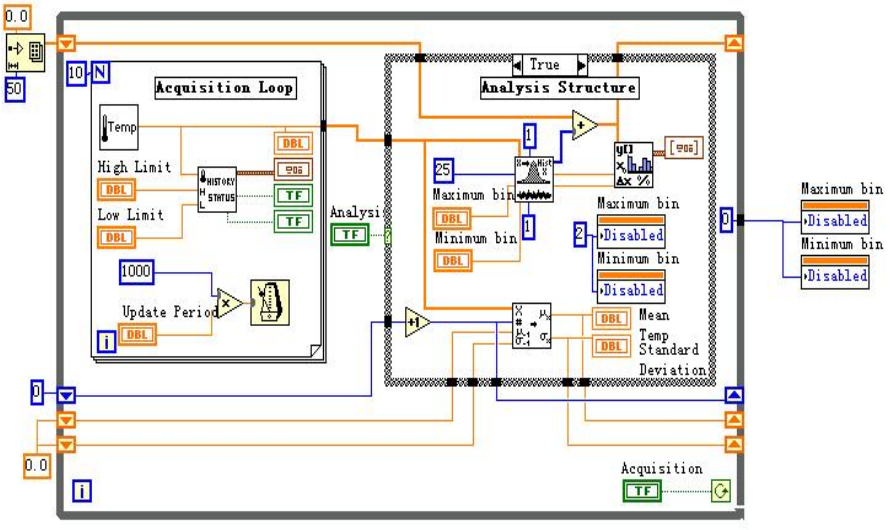


Fig. 1. One example of rare panel of LabVIEW

We are actively exploring ways to use current technology, including computers and the friendly human-computer interface, to aid in instruction and using virtual instrument in an easy way. To this end we are developing a virtual instrument developing system where people can be exposed to easy way to understanding virtual instrument and methods of data collection and processing which they otherwise might not clear because of limited experience and not the expert in software/hardware design in virtual instrument. Herein we put forward a new way to study and use virtual instrument via a system called module instrument(MI), people can use the system to learn how to set-up and operate the equipment and work with the actual device. This system relieves people from the traditional way by which people who intent to a virtual instrument must study the principle of how to design it and understand the working theory of this type of virtual instrument. The system's easy-to-use interface are hence designed and run experiments for people using virtual instrument without experiences rather than for training the users as professional software designers.

1.1 Concept of Virtual Instrument

The concept of virtual instrument is put forward by national instrument corporation: "a virtual instrument is a layer of software and/or hardware added to a general-purpose computer in such a fashion that users can interact with the computer as though it were a traditional electronic instrument." The computer screen presents a

rendition of the control panel of the instrument with controls. The user manipulates the controls with a mouse. The instrument displays readings from signals acquired through a hardware interface to the external world. Commonly, the accepted scheme of virtual instrument is shown as fig. 2

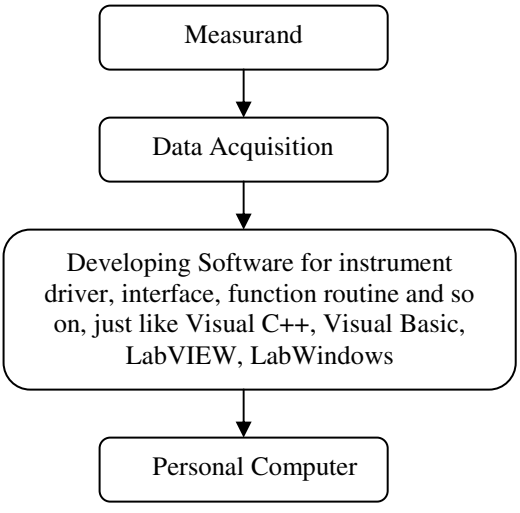


Fig. 2. Scheme of Virtual instrument

1.2 New Way to Virtual Instrument

Instead of spending a lot of time to understand the programming grammar and the development environment just learning how to use the software to control and operate the instrument and to study the working theory of virtual instrument, any individual who just wants to find a simple way to use simple virtual instrument's function, such as data acquisition, may utilize this new way to develop virtual instrument without hardware or software experience needed in other virtual instrument developing environment. With this new tool, people do not acquire necessary experience to debug and program virtual instrument since they can be quickly adept at handling the software to control the actual instrument.

To this end, one new virtual instrument developing environment MI was put forward.

2 Module Instrument

MI is a virtual instrument developing environment (tool), it is devised a module-based framework, designed to allow user who without any previous experience and training to generate virtual instrument easily and naturally, to learn the basic principle of virtual instrument. In order to provide easy understandable interfaces for inexperienced people rather than specialist, MI employs function tightly encapsulated.

The primary consideration of MI is how to encapsulate some common used functions as modules for construct VI, and then how to realize VI via only few required module selection.

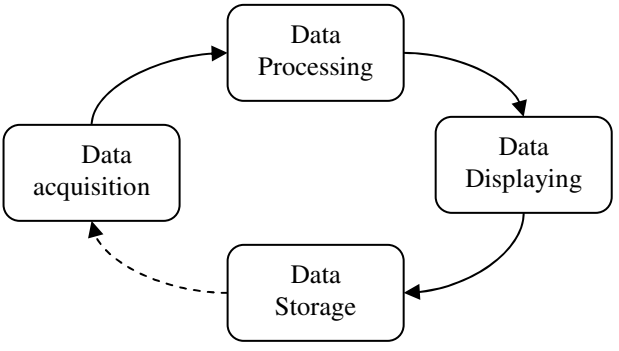


Fig. 3. The common data flow sequence between different modules: the dotted line means that the data flow cycle may be one time or more than one

To encapsulate function as modules, all the function encapsulated can be classified into four types of module: acquisition, displaying, processing and storage. The classification of modules is based on the common data flow sequence when VI works (show as fig.3). All those modules are shown as icons on the toolbar under the environment of module instrument. When user wants to realize one type of VI, he/she just selects corresponding module form those four types modules, and the MI system can automatically connect the selected modules to generate the virtual instrument.

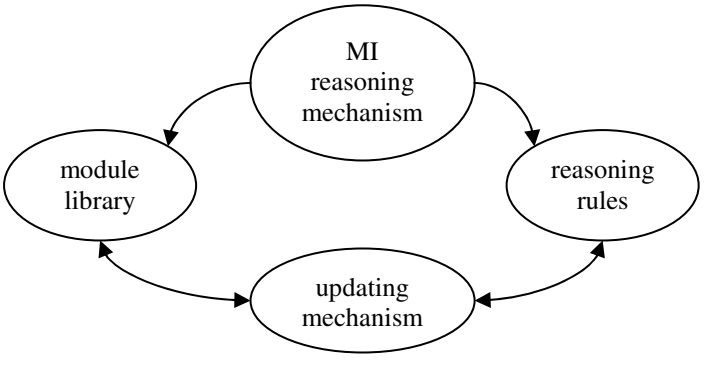


Fig. 4. Relationship scheme of four components for MI system

For the purpose of generate VI based on the module selection, MI system consists four main components: MI reasoning mechanism, module library, reasoning rules and updating mechanism. The relation among those four components is shown as fig. 4: 1)

MI reasoning mechanism is the most important part for MI, this component joints selecting modules to generate VI according to the reasoning rules and is responsible for the coordinating among different modules. 2) Module library is the collection of four types of modules mentioned before and presented as toolbar in MI. 3) Reasoning rules are the basement for MI reasoning mechanism, MI reasoning mechanism uses those rules to determine whether using the selected modules to generate VI is feasible or not, and if feasible, determines the data flow sequence among selected modules.

3 MI and Requirement-Driven

The main aim of MI is simple and easy to use, so how to expose the MI modules to user in an obvious way should be taken into account, requirement-driven [9] is just the idea to tackle this question. Requirement-driven refers the generation of VI is based on the user's most directly requirement which is the foundation of VI generation and is the information source from the user. The main point of requirement-driven for VI generation is relieving user both form complex programming grammar studying and from the deeply understanding instrument inner structure, and the only thing users have to do is putting forward the functional requirement. The idea of requirement-driven for MI is described in fig. 5.

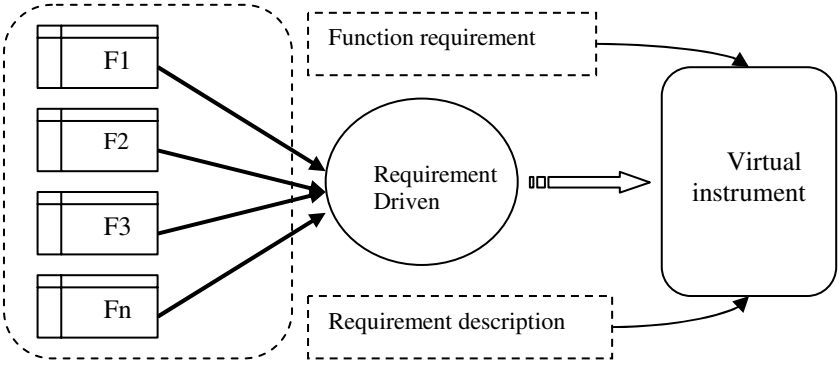


Fig. 5. The exploded map for requirement-driven. Letter F represents function description.

From inexperience users' viewpoint, the simpler one is better. To realize the requirement-driven in MI, the fulfillment of requirement description is just the several times mouse clicks. Based on the highly encapsulated modules, four types of modules are designed as four toolbar user interface in MI and each module is shown as icon on toolbar with the help tip when mouse moves on it. By this way, the use can select one module from those four toolbar respectively; the MI will generate a VI or give selection errors according to the default reasoning rules. Fig.6 shows a data generator from acquisition toolbar in which all available acquisition modules lie.

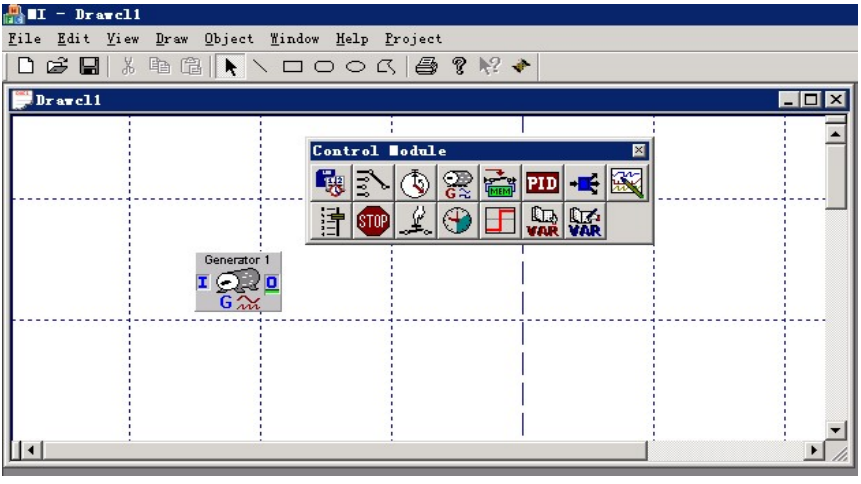


Fig. 6. The selection operation interface for user

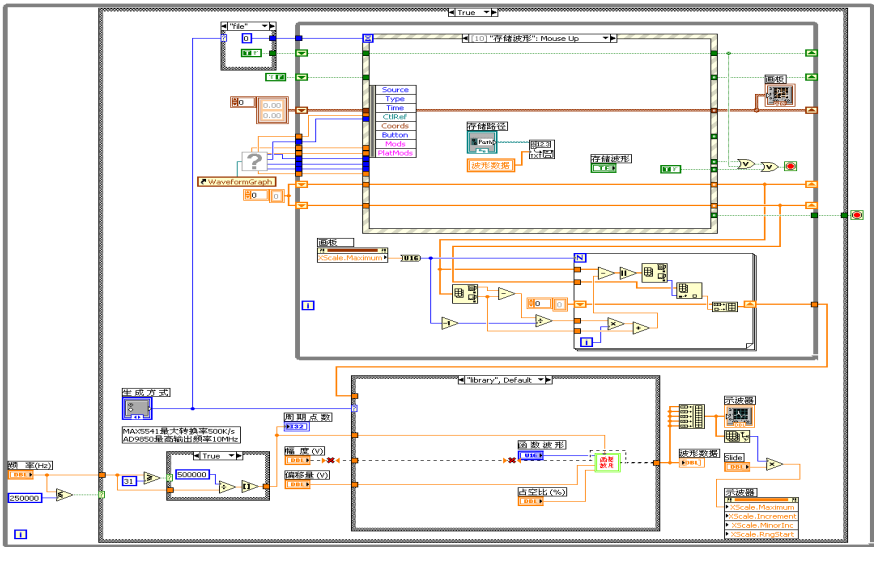


Fig. 7. Screen shot of rare panel for virtual digital oscilloscope under LabVIEW environment, the line connect different type of icon is the idea of data flow. Different type graphic means different functional units.

4 Example of Virtual Digital Oscilloscope

In this section, we present two examples with the same function: virtual digital oscilloscope which displays the simulated waveform. One example of virtual digital oscilloscope is designed by LabVIEW, and the other by MI. Through those two

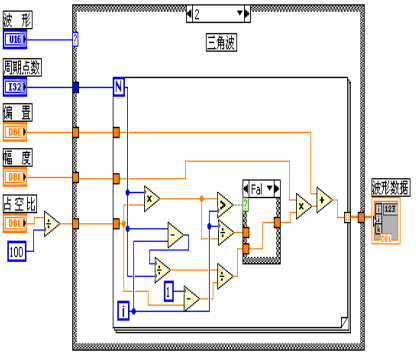


Fig. 8. The rare panel screenshot of design for triangular wave

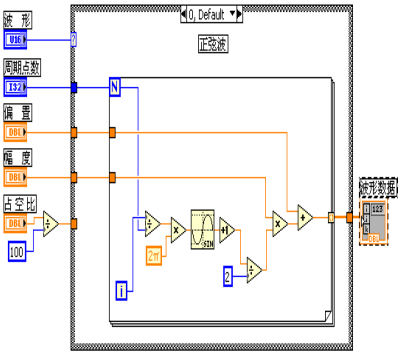


Fig. 9. The rare panel screenshot of design for sine wave

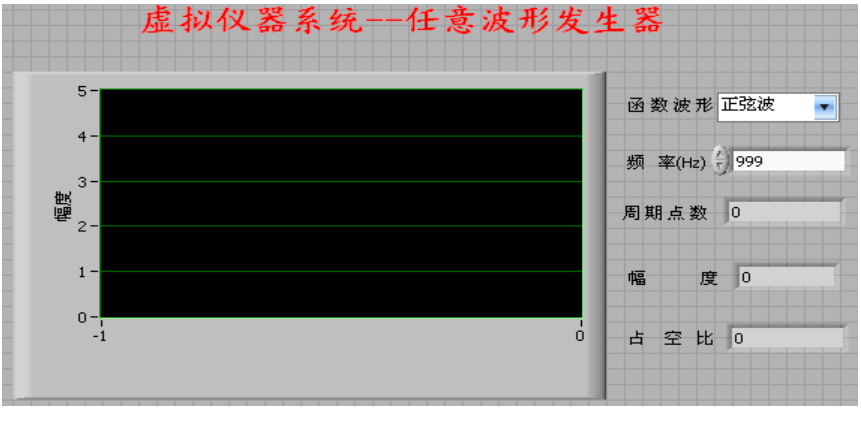


Fig. 10. The front panel screenshot, which is also the user interface of VI

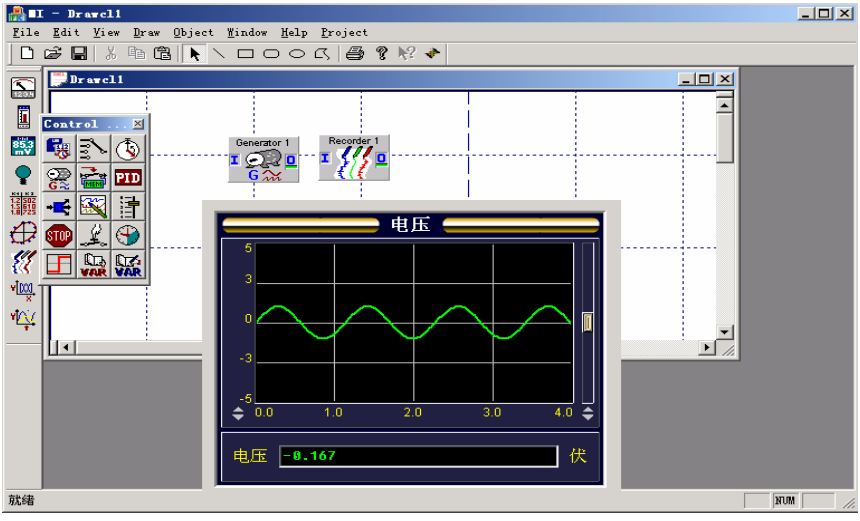


Fig. 11. The screenshot of virtual digital oscilloscope designed by MI

examples, we can get the general impression of MI's objective and user- group faced. The comparison between MI and LabVIEW means to explain why laypeople can use MI easily realizes common functional VI. Fig.7 is the rare design panel of virtual digital oscilloscope under LabVIEW environment; Fig.8 and 9 are the detail design of triangular wave and sine wave; Fig. 10 is the interface of this VI under LabVIEW. Fig.11 is the interface and VI generation under MI environment. From the comparison, MI provides a more easy-to-use interface to realize virtual digital oscilloscope than LabVIEW and gives inexperienced people a more simple way to realize VI and know the VI's working flow from the two type of module selection (one module is the data generation module from acquisition toolbar and the other is the data display module from displaying toolbar). The generation of virtual digital oscilloscope only needs two modules and just two time mouse clicks.

5 Conclusion

In this paper, we introduced a virtual instrument developing environment called MI which exploits a novel reach to virtual instrument using and learning for inexperienced people in related field. The long series of experiment test and evaluation reports from inexperienced people and people who are familiar with LabVIEW encourage us to access that MI is possible to transfer to popularity among untrained people. Unfortunately, the generation of MI is limited. The VI from MI can be the replacement of common used traditional instrument and can not be used to realize very complex functional virtual instrument. Also, there has been no testing of the usability of cross-platform; MI is just can be run on windows operation system. Much more study of this area is sorely needed. In addition, there needs to be a study of how MI reasoning mechanism can apply to generation VI based on many modules selection. Current work has been limited to a few modules, MI reasoning mechanism

can work well if the selected modules are no more than two in each type. In order to make MI simpler, much more technique details about human and computer interaction must be considered under MI. Future work will involve the develop of a prototype with a more intelligent user interface.

References

1. Wong, W.: LabVIEW embraces graphical object-oriented programming. Electronic Design, vol. 54(21), pp. 38–40. Penton Publishing Co, Cleveland, OH 44114-1503, United States (2006)
2. Leping, Y., Haitao, L.: The application of virtual instrumentation in aerospace engineering education, Shanghai (2001)
3. Anon. LabVIEW 20th Anniversary edition emphasizes communications test. Electronic Design 54(18), 33–34 (2006)
4. fan, Y, Lin, J.: A Wireless LAN-Based Robust and Scalable Virtual Laboratory for E-learning. In: Pan, Z., Aylett, R., Diener, H., Jin, X., Göbel, S., Li, L. (eds.) Technologies for E-Learning and Digital Entertainment. LNCS, vol. 3942, pp. 322–325. Springer, Heidelberg (2006)
5. Elliott, C., Vijayakumar, V., et al.: National Instruments LabVIEW: A Programming Environment for Laboratory Automation and Measurement. JALA - Journal of the Association for Laboratory Automation 12(1), 17–24 (2007)
6. Navaee, S.: Computing and programming with LabVIEW, Salt Lake City, UT, United States, American Society for Engineering, Washington, DC 20036, United States (2004)
7. See, A.: Hands-on learning and implementing using LabVIEW for undergraduates in 13 weeks, Salt Lake City, UT, United States, American Society for Engineering, Washington, DC 20036, United States (2004)
8. Sevgi, L., Uluisik, C.: A labview-based virtual instrument for engineering education: A numerical fourier transform tool. Turkish Journal of Electrical Engineering and Computer Sciences 14(1), 129–152 (2006)
9. Yong-kai, F., Jun, L., Tian-ze, S.: Program generation mechanism based on module relationship match reasoning. Journal of Jilin University(engineering and technology edition) 36(6), 939–944 (2006)

Agents in m-Learning Systems Based on Intelligent Tutoring

Vlado Glavinic[1], Marko Rosic[2], and Marija Zelic[2]

[1] Faculty of Electrical Engineering and Computing, University of Zagreb, Unska 3, HR-10000 Zagreb, Croatia
[2] Faculty of Natural Sciences, Mathematics and Kinesiology, University of Split, Teslina 12, HR-21000 Split, Croatia
vlado.glavinic@fer.hr,
{marko.rosic,marija.zelic}@pmfst.hr

Abstract. Intelligent tutoring systems (ITSs) represent a particular kind of e-learning systems, which base their operation on the simulation of a human teacher in the learning and teaching process. With the advent of the mobile computing paradigm, m-learning systems, as the "portable and personal" fashion of e-learning, paved the way to the introduction of mobile intelligent tutoring. Mobile intelligent tutoring systems (MITSs) are targeted to fit into a mobile learner's daily routine without disrupting her/his other activities, but conversely enhancing the efficiency and effectiveness of learning in the context of handheld terminals of restricted capabilities. As in the non-portable ITS counterparts, MITSs' tasks are taken over by agents, making them agent-based systems. In this paper we discuss the mobile intelligent tutoring paradigm, as well as the agent types to be used in the m-learning environment along with the presently affordable agent infrastructure enabling MITS implementation, and corroborate this with the description of a mobile intelligent tutoring model we are developing.

Keywords: agents, intelligent tutoring systems, m-learning, agent-based systems, mobile intelligent tutoring systems.

1 Introduction

Information and communication technology (ICT) has become an integral part of the educational system for either assisting a teacher in realizing a traditional lecture or in completely substituting the teacher. *E-learning* took place between ICT on the one hand and education on the other. The American Society for Training and Development defines e-learning as "instructional content or learning experience delivered or enabled by electronic technology" [1]. A subclass of e-learning – *mobile learning* (*m-learning*) – is nowadays attracting a lot of attention as the next kind of learning, the respective implementation issues however still being the object of more detailed further elaboration. Obviously, the development of this learning paradigm should be considered mostly through developments in a (possibly high-bandwidth) wireless infrastructure represented by technologies such as GSM, UMTS or WiMax. By using the above infrastructure, it is possible to perform networking of mobile devices

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 578–587, 2007.
© Springer-Verlag Berlin Heidelberg 2007

having a number of characteristics interesting to the learning process: (i) small size and high portability, (ii) flexibility in supporting a wide range of learning activities and (iii) a relatively cheap cost of technology. In fact, we are now at the break of a new world of ubiquitous computing that will bridge the gap between our largely sedentary lives and new forms of nomadic life. Hence developments in both wireless data communication and mobile devices facilitate the emergence of a trend of nomadic computing, and consequently of a new fashion in exercising the learning and teaching process. Within the educational context these new nomads will be impersonated by students, teachers, employees, businessmen and customers, who will be traversing both digital and physical landscapes to gather resources, enlarge their knowledge and test their skills, conduct business, and will overall experience both state-of-the-art technologies and services geared to new nomadic lifestyles.

Mobile learners usually work on their own, without external expert support, and possess disparate learning backgrounds. With little time on hand for educational activities and no human expert available for immediate help, these learners require the least learning "overhead" possible, in order to be able to efficiently focus on their actual learning tasks. This in turn requires m-learning systems capable to adapt their behavior to learners' individual requirements [2]. Within such a framework the opportunity to use *intelligent tutoring systems* (ITSs) is highly welcome. As it is already known, ITSs are e-learning systems which base their work on the simulation of a human teacher in the learning and teaching process [3]. Hence, *mobile intelligent tutoring systems* (MITSs) are the e-learning systems used in the m-learning context [4], which provide intelligent support to the learners. Such systems are intended to fit in the mobile learner's daily routine without disrupting her/his other activities, she/he being able to learn while on a bus, while waiting for a friend in a restaurant or some other appointment, between lectures and the like. Advantages of using MITSs include adaptation to both learner's knowledge and abilities as in conventional ITSs, as well as to her/his location and needs in that location [5].

Another question to be considered when adopting the m-learning paradigm is the adaptability of data services, which should provide for such issues as the display of high-quality images on PDAs. This obviously calls for a new technology, which will be based upon a *middleware* that will support both development and runtime execution of *distributed agent-based systems* to perform adaptability of data services. This middleware should anyway enable the systems to work both in the wired and wireless environment.

According to the previous exposition, MITSs are agent-based systems. Agents met within MITSs conceptually belong to the personal agents' type. According to Case [6] *personal agents* are agents intended to satisfy users' demands and possessing characteristics of proactiveness and adaptability, with the main purpose to help a user by performing actions on her/his behalf. Whether a personal agent will adequately decide in which moment to perform a certain action depends on the knowledge it possesses of the user. Hence the way a personal agent registers information on its user is of key importance. It initially has no knowledge on both user's needs and habits, but starts with a predefined user's profile created for an average user, and only afterwards acquires this knowledge by using some of the following methods [7], [8]: (i) registering the way a user performs actions, (ii) directly from the user's instructions, (iii) from the user's examples and (iv) from the other personal agents.

The paper is structured as follows. Section 2 contains an overview of intelligent tutoring systems; Section 3 presents agent tasks in mobile intelligent tutoring, an agent-based mobile intelligent tutoring model as well as the respective development infrastructure, while Section 4 concludes the paper.

2 Intelligent Tutoring Systems

Intelligent tutoring systems are a generation of systems intended for support and improvement of the learning and teaching process. Their work is based on simulation of a human teacher in the learning and teaching process. A question should be immediately raised about the meaning of the phrase "simulation of a human teacher in the learning and teaching process" as well as which characteristics of a human teacher a tutoring system should possess to be attributed as *intelligent*. The teaching paradigm of an intelligent tutoring system is based on the following features of a human teacher [3], [9], [10], [11]: (i) possession of domain knowledge, (ii) possession of knowledge on teaching methods, (iii) capability of diagnosing students' knowledge, (iv) capability of comparing the level of students' knowledge with the domain knowledge i.e. determining the level of students' domain knowledge acquisition and (v) capability of adjusting the teaching methods to the student.

A teacher in traditional educational institutions generally does not have the possibility to adjust her/his teaching methods to individual student's needs and mostly adjusts them to a student group. Conversely, ITSs adjust themselves to individual student's needs, and should thus contain knowledge on the subject being taught, on the way of teaching and on techniques of diagnosing students' knowledge. By adjusting itself to individual students' needs, the learning and teaching process is approaching the ideal of assigning each student her/his own teacher (1-1 relationship), resulting in significantly better results with respect to the traditional 1-n educational process. According to Major [12], the adjustment to individual students' needs is performed by an ITS in the following ways: (i) varying the difficulty level of teaching contents presented to a student, (ii) adjusting the presentation manner, (iii) providing help to the student depending on his results and behavior during system use, and (iv) adjusting the processing order of individual teaching units.

The four main interconnected modules met in the architecture of a typical ITS include (i) a domain knowledge module with the domain knowledge base, (ii) a teacher module for guiding the learning and teaching process, (iii) a student module with information specific for each individual student, and (iv) a suitable user interface module supporting interaction among the student, the teacher and the domain knowledge [3], [13], [14].

The *domain knowledge module* contains knowledge created by an expert in the domain. Separating the expert knowledge from the rest of the system is one of the prerequisites of achieving reuse of that knowledge. Reusing such knowledge is an important element that lowers costs of developing an ITS.

The *student module* creates the student's model. The student's model contains data on the student's knowledge, her/his assumed capabilities, foreknowledge, and behavior while learning as well as the teaching methods that suite the student best. Using the student's model is necessary in adjusting the system to individual students' needs.

The *teacher module* represents the control part of the system since it guides the learning and teaching process using information gained from the other modules. This module contains knowledge necessary for making decisions about the teaching strategy: it uses information on the current student's knowledge and, based on comparison with the expert knowledge, performs the necessary corrections of the teaching process. The corrections refer to (i) the method of teaching contents presentation, (ii) the adjustment to the student's characteristics of the level of difficulty of the teaching contents, as well as (iii) recommending subsequent actions to the student.

The *user interface* module contains mechanisms used both by the student and the ITS to intercommunicate, possibly including methods of adjusting the interface to a student's individual needs. As the student experiences the system exactly through this module, its implementation quality influences teaching quality. Recording the student's activity during her/his interaction with the user interface enables inferring conclusions on interaction styles that suite the student best. Based on these data it is possible to adapt the user interface to the student.

3 Mobile Intelligent Tutoring

3.1 Agents in Mobile Intelligent Tutoring

Transfer from classical intelligent tutoring to mobile intelligent one can be achieved by assigning the tasks of classical ITSs to the personal agents. Such a transfer into the m-learning environment is natural because each module of an ITS can be regarded as a personal agent. The tasks which the personal agents can take over are enumerated below, and worked out in some detail in the following:

a. taking over the role of a human teacher,
b. adapting the user interface to individual user's needs,
c. enabling cooperative learning,
d. enabling competition,
e. searching the information space,
f. help in using virtual reality systems,
g. adjusting the system to the device used for accessing the system.

3.1.1 Taking over the Role of a Human Teacher

According to King [15] an agent is a computer substitute of a person or a process that fulfills a specific purpose or activity. This entity is capable of making decisions that are similar to human ones. The agent has enough insight into personality of the person it substitutes so it can make decisions on her/his behalf. Following this definition the agent should be able to adapt itself to different educational situations like a human teacher.

3.1.2 Adapting the User Interface to Individual User's Needs

The present paradigm of using a computer is based on direct manipulation of the interface elements [16] where a user gives instructions by using the graphical user interface. Adaptation of the user interface on the simplest level enables the user to specify

different interface parameters based on personal preferences. More complex adaptations call for user interface agents, which are able to learn the user's habits while using the system and subsequently to suggest automatic completion of recognized action sequences.

3.1.3 Enabling Cooperative Learning

Cooperative learning is one of the common learning paradigms present e-/m-learning systems are based upon. The cooperative learning service enables a student-student interaction aimed at solving a mutual problem. With such learning style, established groups of users get tasks to be solved by joint work. Agents within systems based on such a learning paradigm can be used in forming cooperative learning groups. Students' personal profiles provide information necessary to establish groups of users with similar capabilities and interests [17].

3.1.4 Enabling Competition

Competition is an especially important learning component in areas where the student, who has completed the learning process, is exposed to competition, e.g. buying/selling stock shares or carrying out war operations. Competition can be carried out by placing human participants on rival sides. As such type of competition often shows a negative influence on the student's independence of the system access time, agents are usually engaged as rivals to human competitors.

3.1.5 Searching the Information Space

The present information infrastructure enables access to vast volumes of data. M-learning systems can use these data for teaching a distant student. The systems enable access to different knowledge sources via a unique user interface (mostly a Web browser's one). The student having access to different knowledge sources is not spared the problems that users searching the information space generally face nowadays. The user's access to different information sources may be complicated due to distribution of data and their heterogeneousness or information overload.

It is therefore evident that application of personal agents for searching the information space would enable easier performing of complex actions which the user carries out to obtain the desired results. Having insight into the user's personal profile, the personal agent can search the information space instead of the user, selecting data it finds to be of interest to the user.

3.1.6 Help in Using Virtual Reality Systems

Virtual reality (VR) systems use simulations to familiarize students with conditions in the real environment, where they will work after acquiring required knowledge and skills. These systems are especially useful in such areas where learning and training in the real environment is either too expensive or too dangerous. Examples of such areas are flying an aircraft, mastering of surgical skills or attaining knowledge necessary for maintenance, surveillance and intervention within expensive technical systems.

In VR based educational systems, students should be provided with a "helper" that will guide them in using the system. This function can certainly be assigned to an agent, which should follow the student through the VR space and when necessary suggest her/him the appropriate activities. Furthermore the agent will be recording the student's

activities in order to enable the analysis of the student's use of the system. In case the student uses a networked VR system, her/his agent should be capable of communicating with other students' agents so as to accomplish possible common tasks.

3.1.7 Adjusting the System to the Device Used for Accessing the System

By using agents it is also possible to achieve adaptation to the device used for accessing an m-learning system's resources. Such an adaptation is necessary since all devices do not have equal capabilities. This primarily means different display capabilities, amount of memory and processing power. The student's personal agent should adapt content distributed to the student taking all this into consideration. It is also necessary to take into consideration that the student will not always access the system by using the same device.

An example of the infrastructure enabling the implementation of the above described model is illustrated in the following.

3.2 Example Agent Infrastructure

A candidate agent infrastructure which should support the development of an m-learning system should obviously comply with quite a number of requirements, but the paramount one is certainly its conformance to some open standard. The open source package JADE (Java Agent DEvelopment Framework) is a middleware for development and runtime execution of distributed agent-based systems [18], which complies with FIPA (The Foundation for Intelligent Physical Agents) standardization.

JADE runtime environment instances are *containers* (since they can contain several agents), and a set of all containers forms a *platform* providing a homogenous layer above a variety of hardware, operating systems, network types etc. This means that the agent platform can be distributed across machines which do not even have to share the same operating system. JADE agents can transparently communicate across containers and even platforms.

The agent platform dynamically evolves with agents appearing and disappearing based on needs and requirements of the application. Each agent is a peer capable of initiating communication, providing services, searching and discovering other agents, entering and leaving the platform anytime.

In order to properly address the resources of mobile devices, which are limited in terms of memory and processing power, as well as the characteristics of wireless links

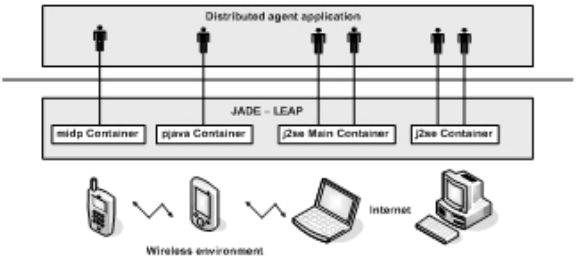


Fig. 1. JADE-LEAP runtime environment

such as high latency, low bandwidth and intermittent connectivity, the LEAP (Lightweight Extensible Agent Platform) add-on was additionally created. It replaces some parts of the JADE kernel, forming a modified runtime environment called JADE-LEAP ("JADE powered by LEAP") which enables deployment of agents on lightweight devices. Figure 1 shows an example of the JADE-LEAP runtime environment.

3.3 Agent-Based Mobile Intelligent Tutoring Model

Most ITSs are designed for use on standard desktop computers or laptops and are not customized for eminently mobile devices such as mobile phones or PDAs. In developing a framework for such systems to fit into the mobile environment with maximum efficiency and minimum effort the main intent is not to reinvent the ITS but rather to insure the integration of the existing "wireline" and the new wireless (GSM, UMTS, WLAN, various satellite systems) environment. However, even though an MITS might seem similar to a conventional ITS, implementation requirements are much more complex. The mobile computing environment is namely very different with respect to the present traditional (non-mobile, i.e. wireline) distributed systems milieu; bandwidth, delay, error rate, interference and the like, may change dramatically as a user moves from one location to another, thus changing the computing environment. Also, the variety of mobile workstations, handheld devices and smart phones is increasing rapidly as well as the range of their capabilities, spanning from low-performance mobile phones to high-performance laptops.

(M)ITS development represents an expensive process, which envisions the engagement of experts in computer science, teaching methodology as well as domain knowledge. Decreasing the respective costs can be achieved by insuring (M)ITS interoperability. An agent-based mobile intelligent tutoring model that enables interoperability of different (M)ITSs by sharing knowledge in the common information space is presented as follows.

Mobile intelligent tutoring is in this model performed by using knowledge in a common information space. The common information space is shared by classical ITSs as well as users' personal agents that take over the ITS's tasks (see Figure 2).

First of all, let us note that sharing knowledge in the common information space may be complicated due to the fact that different systems differently display and store knowledge. Furthermore, different systems use (i) different terminology for the same

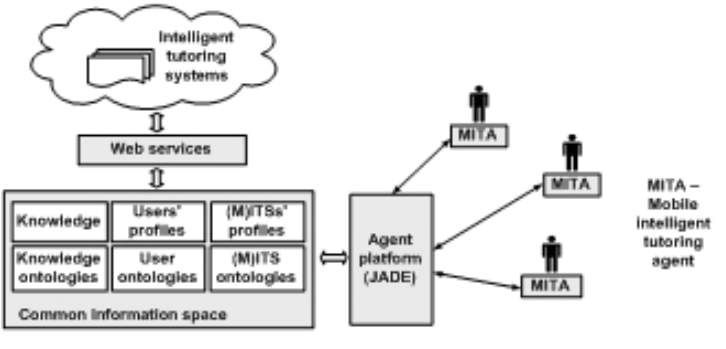


Fig. 2. Agent-based mobile intelligent tutoring model

domain knowledge elements, (ii) the same terminology for different elements, (iii) different ways of connecting these elements as well as (iv) different inference rules on the elements. The common solution to this problem is the use of *ontologies*, which define domain concepts and their attributes, relationships between these concepts as well as inference rules within the domain [19].

The common information space should include three types of ontologies, which are to enable publishing along with sharing knowledge, users' profiles as well as data within (M)ITSs.: (i) knowledge ontologies, (ii) user ontologies and (iii) (M)ITS ontologies. (M)ITSs access these ontologies in the common information space by using adequate Web services (e.g. [20]).

A *mobile intelligent tutoring agent* (MITA) takes over the tasks of classical ITSs in the new environment, see Figure 3. Therefore it should contain knowledge on the subject being taught, on the person being taught and on the way of teaching. However, MITA should additionally contain knowledge on capabilities of available terminal devices in order to adjust domain knowledge presentation styles to capabilities of the particular device used for accessing the system. Furthermore the agent should contain knowledge on available networks and their characteristics so it can adjust itself to possible changes in the environment. MITA enables communication among currently active users as well as connecting users with similar interests. This functionality is supported by sharing users' personal profiles within the common information space. MITA also enables testing of mobile users. While performing these tasks the agent is adjusting itself to users' capabilities, terminal devices and networks.

Such an approach to MITS development has the following advantages compared to MITS development from scratch: (i) teaching contents are being used, which pertain to classical ITSs, thus lowering costs of developing MITSs, (ii) by using MITSs classical ITS users can actually learn anyplace and anytime, and (iii) the group of potential intelligent tutoring users is expanding.

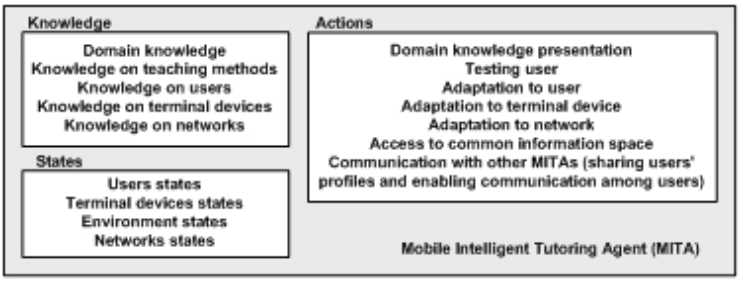


Fig. 3. Mobile intelligent tutoring agent structure

4 Conclusion

M-learning is getting a lot of attention these days as the next kind of learning. We live in an information era where people constantly demand fast and reliable ways of accessing any kind of information, and expect to get this information anytime and anywhere, by using their computers, laptops, mobile phones or whatever technology they

have at hand. However, the m-learning environment represents a challenge for application developers. Firstly, a user can have more networks available at a time, each one with its own characteristics regarding bandwidth, delay, error rate, interference and so on. Secondly, the user can access information by using various mobile or stationary devices, with different processing power, amount of memory and display qualities. A promising approach to solving the above mentioned problems is the use of agent-based technology. Agents' characteristics such as autonomy, learning capabilities, proactiveness and social skills enable adaptation of m-learning systems to the mobile environment.

This paper presents an agent-based mobile intelligent tutoring model that is being currently developed. The proposed model simplifies development of MITSs by reusing knowledge within classical ITSs. The tasks of classical ITSs are taken over by agents, taking into consideration the specifics of the m-learning environment. The agents perform adaptation of the system to the network used for accessing the system as well as the terminal device.

Acknowledgements

This paper describes the results of research being carried out within the project 036-0361994-1995 *Universal Middleware Platform for e-Learning Systems* funded by the Ministry of Science, Education and Sports of the Republic of Croatia.

References

1. American Society for Training and Development: A Vision of E-Learning for America's Workforce. Report of the Commission on Technology and Adult Learning (2001)
2. Kinshuk: Adaptive mobile learning technologies. GlobalEducator.com (2003)
3. Stankov, S.: Isomorphic Model of the System as the Basis of Teaching Control Principles in the Intelligent Tutoring System. PhD Dissertation. Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture University of Split, Split (in Croatian) (1997)
4. Goh, T.T., Kinshuk: A discussion on mobile agent based mobile web-based ITS. In: Proc. Int'l Conf. on Computers in Education. Los Alamitos pp. 1514–1515 (2002)
5. Cui, Y., Bull, S.: Context and learner modeling for the mobile foreign language learner. System 33, 353–367 (2005)
6. Case, S., Azarmi, N., Thint, M., Ohtani, T.: Enhancing E-Communities with Agent-Based Systems. Computer, pp. 64–69 (2001)
7. Maes, P.: Agents that Reduce Work and Information Overload. Comm. ACM 37, 30–40 (1994)
8. Kuflik, T., Shoval, P.: User profile generation for intelligent information agents. In: Proc. 2nd Int'l Workshop on Agent-Oriented Information Systems. Stockholm, pp. 63–72 (2000)
9. Patel, A., Kinshuk, R.D., Oppermann, R.: Intelligent Tutoring Systems: Confluence of Information Science and Cognitive Science. In: Proc. Informing Science Conference. Krakow (2001) pp. 392–397 (2001)

10. Castro, E., Cuadra, D., Martinez, P., Iglesias, A.: Integrating Intelligent Methodological and Tutoring assistance in a CASE platform: The PANDORA experience. In: Proc. Informing Science & IT Education Conference. Cork (2002)
11. Ainsworth, S., Clarke, S.E., Gaizauskas, D.: Using edit distance algorithms to compare alternative approaches to ITS authoring. In: Cerri, A., Gouarderes, G., Paraguacu, F. (eds.) Intelligent Tutoring Systems, pp. 873–882. Springer, Heidelberg (2002)
12. Major, N.: How Generic can Authoring Shells Become? In: Proc. AI-ED-95 Workshop on Authoring Shells for Intelligent Tutoring Systems. Washington, pp. 75–82 (1995)
13. Burns, H.L., Capps, C.G.: Foundations of Intelligent Tutoring Systems: An Introduction. In: Polson, M.C., Richardson, J.J. (eds.) Foundations of Intelligent Tutoring Systems, pp. 1–18. Lawrence Erlbaum Associates Publishers, London (1988)
14. Woolf, B.P.: AI in Education. In: Shapiro, I., Stuart, C. (eds.) Artificial Intelligence - Encyclopedias, pp. 434–444. John Willy & Sons Inc. New York (1992)
15. King, J.A.: Intelligent Agents: Bringing Good Things to Life. AI Expert, pp. 17–19 (1995)
16. Schneiderman, B.: Designing the User Interface: Strategies for Effective Human-Computer Interaction. Addison-Wesley Publishing Company, Reading (1987)
17. Jaques, P., Andrade, A., Jung, J., Bordini, R., Vicari, R.: Using pedagogical agents to support collaborative distance learning. In: Proc. Int'l Conf. on Computer Support for Collaborative Learning. Boulder, pp. 546–547 (2002)
18. Bellifemine, F., Poggi, A., Rimassa, G.: JADE: a FIPA 2000, compliant agent development environment. In: Proc. 5th Int'l Conf. on Autonomous Agents. Montreal pp. 216–217 (2001)
19. Poli, R.: Ontology for Knowledge Organization. In: Green, R. (eds.) Knowledge organization and change. Indeks, Frankfurt, pp. 313–319 (1996)
20. Erl, T.: Service-Oriented Architecture. A Field Guide to Integrating XML and Web Services. Prentice-Hall PTR, Upper Saddle River, NJ (2004)

Digital Game-Based Learning: An Agent Approach

Christian Anthony L. Go and Won-Hyung Lee

Digital Contents Protection and Computer Game Lab,
Graduate School of Advanced Imaging Science, Multimedia and Film
Chung Ang University, Seoul, 156-756 Korea
chipgo@gmail.com,
whlee@cau.ac.kr

Abstract. Digital game-based learning has proven to be a useful and cost-effective alternative to the traditional classroom-based experience. However, current digital learning methods for young learners fail to engage audiences accustomed to interactive media. Moreover, most edutainment games do not offer players a situated learning experience, and those few that do, do not leverage the immensely popular online game market. This paper introduces a Belief Desire Intention (BDI) agent architecture for an online game Non-Player Character that encourages and stimulates situational learning in an online Role-Playing Game.

Keywords: Digital Game-Based Learning, Intelligent Agent, BDI Agent.

1 Introduction

Interactive digital media, specifically video games, have emerged as a dominant new economic, cultural and, recently, educational force. Most people see today's games and assume they are primarily about high-tech graphics and animation. In truth, games are more powerful, they provide situated experiences wherein players are immersed in virtual worlds and are charged with solving complex tasks.

The prevalence of video games in society today has actually re-wired our brains and has made traditional learning less effective. [1] According to the neurology and psychology research in cited in [1], the brain rewires and reorganizes itself according to cultural stimuli. Educators typically view educational gaming as useful mainly for drill and practice, but as gaming environments become more complex, edugames may be more useful in providing alternative paths to concepts hard to reach using traditional methods.

1.1 Online Games

Today's most popular games are mostly online, with the advent of the Massively Multiplayer Online Game (MMOG) genre, majority of which is dominated by the MMORPG sub genre. MMORPGs function as a form of new place for informal sociability by providing the players with spaces for interaction and collaboration aside from the home, school and workplace. [2].

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 588–597, 2007.
© Springer-Verlag Berlin Heidelberg 2007

This paper introduces a BDI agent architecture for a Non-Player Character that acts autonomously and realistically, while, at the same time, gives players a situated learning experience. Its ability to realistically model human behavior by achieving a good balance between reactive and deliberative behavior, while at the same time corresponding to the way people explain their actions and reasoning, have made the BDI Architecture a very popular choice for Agent AI implementations.

2 Related Work

Game-based learning has been extensively used by the military. In fact, the US military is the world's largest spender on, and user of, digital game-based learning. The military uses games to train soldiers, pilots, sailors, tank drivers and hovercraft operators to master their expensive and sensitive equipment. And because modern warfare increasingly takes place on airplane, tank or battleship computer screens without the operator ever seeing the enemy except as a symbol or avatar, simulations are surprisingly close to actual combat. [3].

Tsushima and his colleagues at the Department of Digital games at Osaka Electro-Communication University in Japan have developed a New Generation Authoring System specifically for learning contents called NeGAS. The main aim of his project is to develop an authoring system for learning systems.The user of NeGAS can construct learning contents through the use of a mouse because it provides a direct interface to the contents. [5].

Norling and Sonenberg [7] used the BDI architecture to create interactive game characters. This method was illustrated in a project where models of expert players of Quake2 were developed. They successfully demonstrated that a BDI-agent based programming language could clearly and succinctly capture individual differences.

3 Games and Learning

3.1 Game-Based Learning

A new model of e-learning, commonly called digital game-based learning is emerging as an alternative vision for e-learning. [10] This is a wholly different form of digitally mediated experience with roots in entertainment and is emerging as a very powerful form of learning. In recent years, several research projects, organizations, centers, grants, books and studies have emerged, exploring new visions for game-based technologies in learning. [11] Games provide a familiar method of information delivery via a computer or the Internet. Game players routinely spend hundreds, if not thousands of hours mastering complex skills in digital worlds that are time-consuming, challenging and difficult to master. All of this in a highly competitive, rapidly iterating, Darwinian environment, resulting in game mechanics that are highly refined, embodying a wealth of design knowledge. [8].

3.2 Situated Learning in Games

Learning as it normally occurs is a function of the activity, context and culture in which it occurs (i.e., it is situated). This contrasts with most classroom learning

activities which involve knowledge which is abstract and out of context. Social interaction is a critical component of situated learning -- learners become involved in a "community of practice" which embodies certain beliefs and behaviors to be acquired. As the beginner or newcomer moves from the periphery of this community to its center, they become more active and engaged within the culture and hence assume the role of expert or old-timer. Furthermore, situated learning is usually unintentional rather than deliberate. These ideas are what Lave & Wenger call the process of "legitimate peripheral participation." [14][15] Learning, both outside and inside school, advances through collaborative social interaction and the social construction of knowledge." Brown et al. also emphasize the need for a new epistemology for learning -- one that emphasizes active perception over concepts and representation.

Despite the ambivalence of educators regarding the power of games for learning, it is interesting to note that learning within game environments matches up very well with emerging cognitive science research on how people think and learn. [17] Cognitive scientists coming from different traditions have now come to adopt what is called a “situated view of learning”. This proposes that thinking is not a matter of abstract, symbolic representations, but rather rooted in direct experience and concrete contexts. Numerous compelling illustrations of learning based on situated learning are emerging. The most compelling among these are the learning that is naturally occurring in games like Rise of nations, Civilization III, Lineage and Lineage 2. [18].

4 Social Interactions in an MMORPG

Due to the Internet’s capacity for connecting people across time and space, social networks and communities have been formed. The gaps of race and class have likewise been bridged [20] [21] This is especially evident in online games where players are often unaware of each other’s race, age, marital status and nationality.

The online gaming industry continues to grow by leaps and bounds, with over nine million subscribers worldwide. [22] MMORPGs are played heavily, with the average time spend in-game being 20 hours a week, and often with friends and family. [23]

Of vital importance to the MMORPG experience is the realism and immersiveness of the virtual world. These virtual worlds are experienced more like places and the users feel to be part of something they perceive of as a place. The idea and praxis of a virtual life parallel to real life is promoted to the users. [21][24].

5 Belief Desire Intention (BDI) Paradigm

The underlying philosophy of the BDI framework is based upon folk psychology. This describes the way people *think* that they think-this maps closely with how people ordinarily explain their reasoning, as opposed to an objective view of how the brain actually works. This model represents an abstraction of human deliberation based on a theory of rational actions in the human thinking process. This model of practical reasoning was proposed by Bratman et al. in 1987. [25].

In agent Applications, the BDI theory is the most popular and influential. [26]. BDI agents posses the following qualities:

Beliefs-represents the agent's general knowledge about its environment and itself. An agent can store this information using a database, a set of variables or some other data structure.

Desires-describes the agent's goal, a system state that the agent wants to achieve. These desires represent the objectives that the agent aims to accomplish or, more generally, the priorities associated with them

Intention- the course of action that that agent has chosen to achieve that goal, the deliberative state of the agent.

Plan-a recipe for action for achieving a certain goal. This is used to guide the decision-making process of agents so they need not search the entire space of possible solutions.

Given its beliefs about the world, an agent will intend to do what will achieve its desires (goals), all in accordance to a plan. An intention is formed when an agent commits to a plan in order to achieve a goal. The steps in the plan may be singular, atomic actions or they can also be subgoals. Since the agent does not need to commit to a particular plan or subgoal until the last possible moment, the result is an agent with a good balance between deliberative (planning) and reactive behavior. [27] [28].

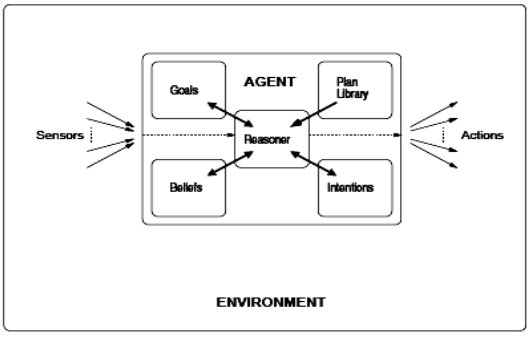


Fig. 1. The Belief Desire Intention (BDI) Architecture

The BDI architecture has been used in numerous dynamic and complex scenarios where agents need to act under incomplete and incorrect information about the environment and other agents.

6 Non Player Character (NPC)

Present implementations of NPCs in games do not simulate realistically believable human interactions Agents in NPC applications must be able to engage players in

interesting and entertaining reactions. The need for this relationship can be best illustrated by the following quote [4]:

"The fact is, in most computer role-playing games, the NPCs don't give any semblance of intelligence. Say you're in a bar and you throw your beer at the bartender one day. The next day you go back, and he's just as happy to see you. That shouldn't happen".

A good NPC must be able to make the game player believe that he/she is interacting with an intelligent, sentient character with goals, beliefs or desires of their own.

7 A BDI NPC

In order to create believable NPCs for games, a BDI agent architecture must satisfy the following conditions:

1. Autonomy
2. Social Ability
3. Reactivity
4. Pro-activeness
5. Persistence

We propose a modular architecture to meet these demands. The main components of the architecture are illustrated in figure 2. Information is collected from the game by agent sensors. These are transformed into a knowledge store to which each of the selected modules has access. This knowledge store represents the NPCs beliefs about its environment.

7.1 Goals

The goal scheduler module is responsible for scheduling of goals. Goals will have to be predefined by the game designer. The goal scheduler updates its schedule based on data from the knowledge store and likewise updates the knowledge store with regards to its progress in goals.

7.2 Behavior

The behavior module is a classical static agent system. It will be responsible for overall agent behavior in the world, walking, running, leaving clues, verifying information, attacking, trading etc. Given data from the knowledge store, it will update its behavior and in doing so will also pass information onto the knowledge store.

7.3 Social Interactions

The social system will be similar to the behavior module. However, we separate this module to add more granularity to our NPC's social interactions with players. This will likewise be a static system but with a wider repertoire of social skills. This will handle social behavior such as talking, arguing, making friends and the like. This

social system will get information and updates from the knowledge store. It also, however, also maintains a social memory of its encounters with people. At the start of a social interaction, this memory is accessed and it is updated at the end of a social encounter. This social memory module will enable our NPC to remember who its allies and enemies are.

7.4 Decision Maker

The decision making module is the selection mechanism that mediates between each system and decides what action to take based on the information it receives from them.

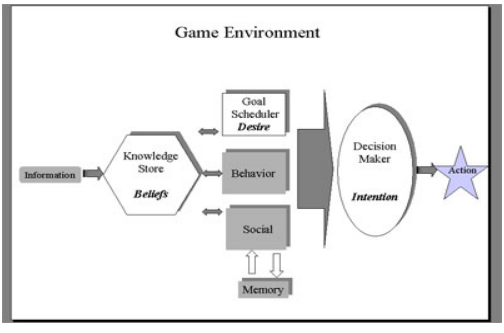


Fig. 2. The NPC BDI Architecture. Note the Separate Social Interaction Module and its dedicated Memory Module.

8 Experiments

8.1 Set-Up

Due to the proprietary nature of most MMORPGs and game companies' reluctance to divulge their source code, we sought an alternative that had all the elements of an MMORPG yet allowed scripting and code modifications. With this in mind, experiments were conducted on the Aurora Game Engine of BioWare- the game engine used for the massively popular Neverwinter Nights RPG franchise, which has been expanded into an MMORPG. [12] We likewise selected this game engine due to the large number of players that regularly interact on it online.

We implemented a prototype NPC using our BDI architecture. The NPC goal was to impart simple yet novel knowledge to the player through its flexible multi-branched dialogue and quest milestones and items. It would verify this information by asking the player a series of follow-up questions at a later time, all in the context of completing a mission. Information was divided into two (2) parts, abstract and contextual information.

Abstract information was given as a set of ten (10) trivial facts that are part of the NPC dialogue but not related to the quest. Contextual information was provided as a total of ten (10) pieces of information, items to be obtained, and player milestones to

achieve. In order for the quest to be completed, the player must have accomplished all required quest milestones, gathered all necessary items, and answered the BDI NPC questions using information gleaned from dialogues with it as well as, other information obtained from ordinary NPCs while traveling the game map.

The average time from the start of the quest to completion is twenty (20) minutes. We designed the game map and arranged the waypoints and destinations with the goals of providing a rich game map and maximizing player travel time. This was necessary to assure that the information retrieved from players has actually been committed to long-term memory and is not simply a rehash from short-term memory.

Fig. 3. Screenshot of a player Interaction with our BDI NPC. The NPC is the man to the right with the moustache.

The test employed ten (10) volunteers, all experienced computer game players whose preferred game is the RPG. We instantiated a single BDI NPC for contextual information and two simple NPCs for abstract information in each test. At different intervals after the testing, the volunteers were given a short quiz regarding the information they had learned from the game. This was conducted at 2, 24 and 48 hours post-experiment.

8.2 Results

As expected, abstract information retrieval was poorer compared to contextual information. On the average, the test subjects correctly recalled 39% of the information for

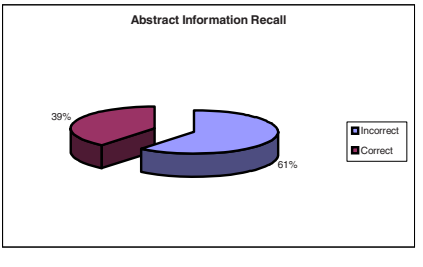


Fig. 4a. Abstract Information Recall

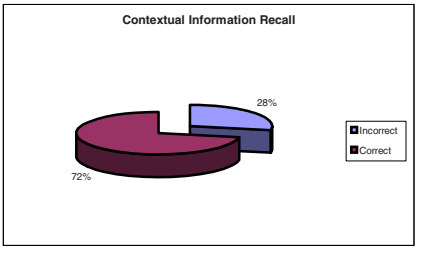


Fig. 4b. Contextual Information Recall

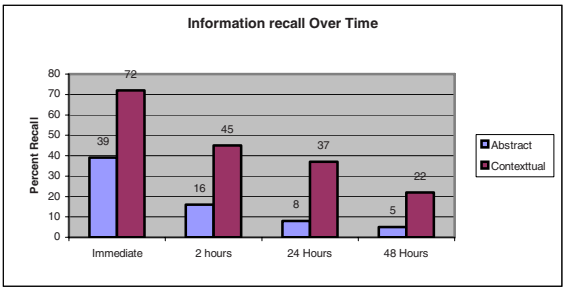


Fig. 5. Percentage of Information recall over time

abstract information. Contextual information recall, however, was much higher, with volunteers correctly recalling 72% of the information. It is interesting to note while the information recalled does slowly degrade over time due to its being only relevant in the game world, the rate of decay for contextual information was considerably less than abstract information. Contextual information recall degraded to 22% from 72% while abstract information fell from 39% to 5% overall after 48 hours.

9 Conclusion

Initial experiments using our prototype has proven that effective, situated digital game-based learning facilitated by a BDI NPC can occur, in the context of an online game, specifically a Massively Multiplayer Role-Playing Game. Aside from contextual data, game players can likewise learn abstract facts, albeit to a lesser extent. While information recall does degrade over time, we attribute this to two factors: the amount of time spent learning (exposure), and, the lack of further reinforcement of what is learned. Despite this, we have demonstrated that an effective and quasi-realistic game BDI NPC can, to some extent, mimic a real player and give game players the social interaction needed for effective situated learning. One of the limitations of our current prototype however, is its high system overhead. To circumvent this, a game designer might make NPCs that are central to a game persistent BDI agents, while relying on traditional NPC techniques to model non-critical characters of the game. This research offers a glimpse into numerous digital game-based learning possibilities for the immensely-large yet still relatively untapped MMORPG market.

9.1 Future Work

Further experiments are currently underway. Immediate work in progress includes refinement of the social module of the BDI NPC to add more realism to player interactions. The present prototype introduces significant system overhead, hence, streamlining the BDI NPC Agent is necessary. Once the final version of our model has been completed, we will commence experiments involving a much larger sample size. This will include game players with different levels of skill and differing game preferences. We will also conduct further reinforcement of lessons learned, and tracking of

how much of the information learned is accurately retrieved over a longer period of time.

Further avenues of research, game company willing, may also include actual implementation on a popular online MMOPG such as lineage and its like.

Acknowledgements. This research was supported by Chung Ang University and Seoul Future Contents Convergence (SFCC) Cluster established by the Seoul R&BD Program.

References

1. Prensky, Marc: Digital Game-Based Learning. McGraw-Hill, New York (2001)
2. Steinkuehler, C., Williams, D.: Where everybody knows your (screen) name: Online games as third places. Journal of Computer-Mediated Communication, 11(4), article 1 (2006)
3. Prensky, Marc: True Believers: Digital Game-Based Learning in the Military, Digital Game-Based Learning. McGraw-Hill, New York (2001)
4. Coco, D.: Creating Intelligent Creatures, Computer Graphics World (July 1997)
5. Tsushima, k.: Progress Toward the Science of game and amusement, Active Media Technology (AMT 2005). In: Proceedings of the 2005 International Conference on (2005)
6. Masuch, M., Rueger, M.: Challenges in Collaborative Game Design Developing Learning Environments for Creating Games, Creating, Connecting and Collaborating through Computing, C5 2005. In: Third International Conference on (2005)
7. Norling, E., Sonenberg, L.: Creating Interactive Characters with BDI Agents. In: Proceedings of the Australian Workshop on Interactive Entertainment IE2004 (2004)
8. Squire, Kurt. Game-Based Learning, Report to the Masie Consortium (2005)
9. Ednar, A.K., Cunningham, D., Duffy, T.M., Perry, D.J.: Theory into practice: How do we link? In: Instructional Technology: Past, Present and Future, Englewood, Co: Libraries Unlimited Inc (1995)
10. Aldrich, C.: Simulations and the future of Learning. Pfeiffer, New York (2004)
11. Games-to-Teach team at MIT. Design Principles of next-generation digital gaming for education, Educational Technology (2003)
12. Neverwinter Nights (NWN), produced by BioWare and published by Infogrames (now Atari)
13. Kirk, J. Belovics, R.: An Intro to Online Training Games http://www.learningcircuits.org/2004/Apr2004/kirk.htm
14. Lave, J.: Cognition in Practice: Mind, mathematics, and culture in everyday life. Cambridge University Press, Cambridge, UK (1988)
15. Lave, J., Wenger, E.: Situated Learning: Legitimate Peripheral Participation. Cambridge University Press, Cambridge, UK (1990)
16. Brown, J.S., Collins, A., Duguid, S.: Situated cognition and the culture of learning. Educational Researcher 18(1), 32–42 (1989)
17. Gee, J.P.: Learning by Design: Games as Learning Machines. (March 2004) http://www.gamasutra.com/gdc2004/features/20040324/gee_01.shtml
18. Gee, J.P.: What videogames have to teach us about learning and literacy. Palgrave Macmillan, New York (2003)
19. Eladhari, M.: Trends in MMOG Development http://www.game-research.com/art_trends_in_mmog.asp

20. Putnam, R.D.: Bowling Alone: The Collapse and Revival of American Community. Simon & Schuster, New York (2000)
21. Go, C.A., Lee, W.H.: Virtual Terrain Evaluation Using Pedestrian Agents. In: Proceedings of the Second International Conference on Design Computing and Cognition (2006)
22. Woodcock, B.: An Analysis of MMOG Subscription Growth-version 18.0. (Retrieved July 17, 2006) (2006) from http://www.mmogchart.com/
23. Yee, N.: The Daedalus Project. (Retrieved July 17, 2006) (2006) from http://www.nickyee.com/ daedalus/archives/001468.php
24. Eladhari, M.: Trends in MMOG Development http://www.game-research.com/art_trends_in_mmog.asp
25. Bratman, M.E.: Intentions Plans and Practical Reason. Harvard University Press, Cambridge, MA (1987)
26. Rao, A.S., Georgeff, M.P.: Modeling rational Agents within a BDI architecture. In: Proceedings of Knowledge Representation and Reasoning, pp. 473–484 (1991)
27. Rao, A.S.: A Unified View of Plans as Recipes, Contemporary Action Theor. Kulver Academic Publishers, The Netherlands (1997)
28. Norling, E., Sonenberg, L.: Creating Interactive Characters with BDI Agents. In: Proceedings of the Australian Workshop on Interactive Entertainment IE2004 (2004)

A Haptic Based Interface to Ease Visually Impaired Pupils' Inclusion in Geometry Lessons

Cédric Gouy-Pailler, Sophie Zijp-Rouzier, Sylvie Vidal, and Denis Chêne

France Télécom Recherche & Développement
28 Chemin du Vieux Chêne, 38240 Meylan, France
cedric.gouypailler@gmail.com,
sophie.zijprouzier@orange-ftgroup.com

Abstract. This paper describes a system based on haptics and sound, to assist in teaching plane geometry to visually impaired pupils. So far the exploration of the geometric figure was performed with an haptic device and construction commands launched with a keyboard. This way of interaction (switching from a device to another) proved to be quite inefficient. This work aims at using exclusively an haptic device to both perceive and build a figure. To do so, command gestures are introduced to construct a figure. Then two questions arise: are command gestures haptically learnable (pupil's hand being haptically guided along the gesture)? Are command gestures usable to solve a task (here a geometry exercise)? Two experiments have been carried out to answer these questions.

Keywords: haptic, learning, gesture, blind, multimodal, geometry, design for all, inclusion, co-design.

1 Introduction

First laws in most European countries have pointed out as the main goal the social inclusion of challenged children into the standard school environment.

Moreover geometry is one of the most difficult subject to teach to blind pupils and one of the most useful at the same time: they need to build a coherent spatial representation of the world and to share with sighted people a set of common point of view (our language is full of geometrical references for example).

So we propose a software (a geometrical figure editor) based on an innovative combination of haptic and audio (speech and non speech) information [5]. Our first prototype offered only keyboard commands. But now interaction is based on 2D command gestures (deictic and semantic). Then two main questions arise: are command gestures haptically learnable (pupil's hand being haptically guided along the gesture)? Are command gestures usable to solve a task (a geometry exercise) with our interface?

In the following section, we draw up a state of the art on 2D gestural commands. Section 3 describes our system called SALOME. Section 4 details the two experiments

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 598–606, 2007.
© Springer-Verlag Berlin Heidelberg 2007

and their results. Sections 5 and 6 present our current and future work at improving the system. Section 7 concludes with perspectives.

2 State of the Art on 2D Command Gestures

In this part, we make a quick overview of various existing 2D gesture interfaces.

First, there are transverse interfaces like StrokeIt[1] or Windows for Pen[2]. In these interfaces, some wrong choices on an ergonomic point of view have been done: for example, the gesture "two parallel lines" for the delete command in Windows for Pen interface is very difficult to do; Z gesture is much easier.

Second there are browser extensions interfaces like Mouse Gestures for Opera [12] and for Mozilla Firefox [11]. We can notice here several drawbacks: for example, in the first one, similar gestures are used for different commands; in the second one, there is no circular gesture whereas our main natural gestures are curve gestures.

Third there are specific interfaces like GRIGRI (Gesture Recognition on Interactive Graphical Radar Image) [8], "Input interface for interaction with in-vehicle system" [9] or DigiStrips (electronic strips prototype for air traffic controllers) [10]. These interfaces, which are dedicated to applications with risks, respect the following interesting guidelines:

- gesture selection criteria are rapidity, discrimination and mnemonic;
- semantic helps gesture understanding and usability;
- graphic similarity is not a sufficient criteria to evaluate similarity between two gestures (a gesture beginning at the right and ending at the left is clearly a different gesture than a gesture beginning at the left and ending at the right, even if those two gestures are a single horizontal line).

In our interface, gesture vocabulary for standard commands (like confirm, validate, next or previous) has been defined thanks to an ergonomic study that focused on the design of a usable interface through force feedback and command gestures for blind users [6]. In a quite innovative way, the authors concentrate on gestures properties only, on their capacity to be easily memorized, recalled and executed, independently from any visual consideration. They explore the motor compatibility (graphomotor constraints and gesture control), cognitive compatibility (gesture mental and mnesic representation) and semantic compatibility (mental representations of the command and of the gesture meaning).

3 System Description

Our system (called SALOME) is a multimodal geometrical figure editor that allows visually impaired pupils to solve secondary school level geometrical construction exercises[3]. SALOME relies on a haptic force feedback device which is used as a pen to write with friction on a virtual haptic page of a geometry notebook.

[1] http://www.tcbmi.com/strokeit/

[2] http://support.microsoft.com/default.aspx?scid=kb;en-us;87821#kb2

[3] This work is partially funded by the European project MICOLE (http://micole.cs.uta.fi:8080/Plone/)

3.1 General Description

This haptic pen is the only interaction device (no keyboard at all). It is both an input and output device, in order to facilitate human computer interactions:

– it is used to perform contextual commands (input),
– it is also used to perceive the geometrical figure (output).

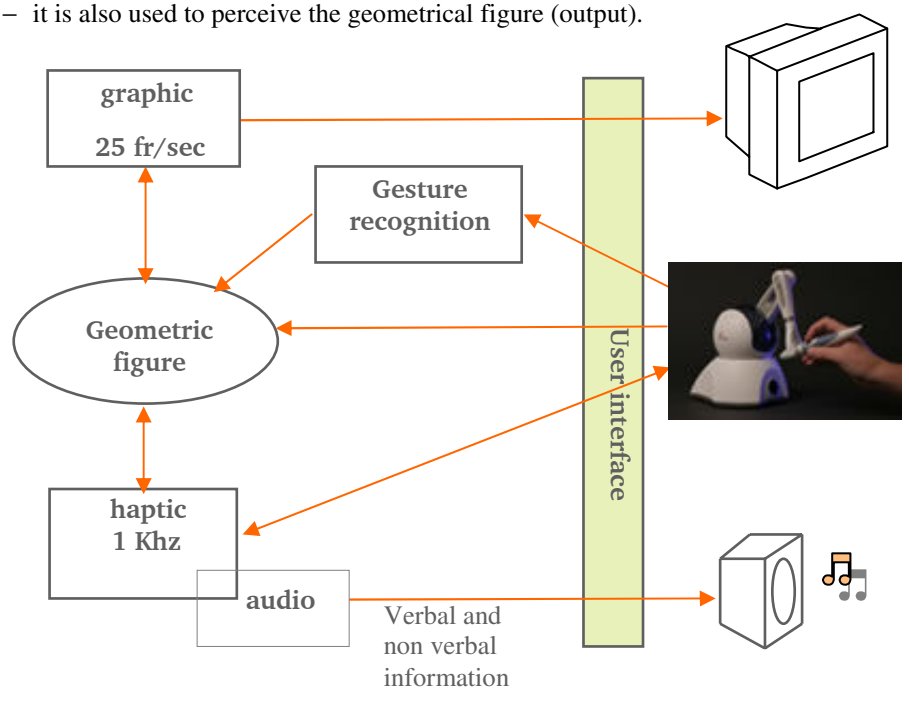


Fig. 1. SALOME, a multimodal application (force feedback, gesture recognition, sound and visual)

3.2 System Functionalities

This system is designed to provide four main functionalities to the visually impaired pupils in order to construct and interact with geometrical figures:

Figure construction. The user can make nice and precise drawings. To do this, he has access to different construction command gestures in order to draw points, segments, lines, triangles, circles for example.
Figure reading. To understand the figure, visually impaired pupils move their *haptic pen* and can feel straight lines, points, curves, that they have previously constructed. Geometric figures are coded as haptic *magnetized rails* that attract the pen towards the different elements of the 2D figure. Each element has also an audio description (synthesizer voice) that enhances the proprioceptive perception of the element. The sequential audio-haptic perception enables to shape some spatial representation of this figure.

Figure measuring. Thanks to command gestures, the user can quantify any distance between two points: a synthetic voice says the length in centimetres.
Search mode. This functionality (launched by a command gesture) is very useful for visually impaired pupils who get often lost in the haptic page while exploring or constructing a figure. Thus, the user can find any geometric element of the figure, his haptic pen being sequentially attracted towards the different elements of the figure.

3.3 Command Gesture Based Interaction

Gestures are simple symbols the user "draws" on the haptic virtual plane using the haptic pen. Command gestures are interpreted thanks to a gesture recognition engine named SYMBAL-latin[4] developed at France Telecom R&D.

In our previous work [3], we used two different horizontal planes: one for gesture recognition and one for figure exploration. However, this kind of interaction has pointed out a major limitation from the ergonomic point of view: the spatial context of figure exploration is lost when switching from figure plane to control plane; thus there is no possibility to contextualize commands and link them to the geometric figure already constructed. For a blind person, contextual awareness is really important so this two-planes solution was finally rejected.

For this reason, we have integrated command gestures using a new interaction mode, as described in figure 2. Now, one single plane is both used for figure exploration and for command gestures. When the user holds down the stylus switch

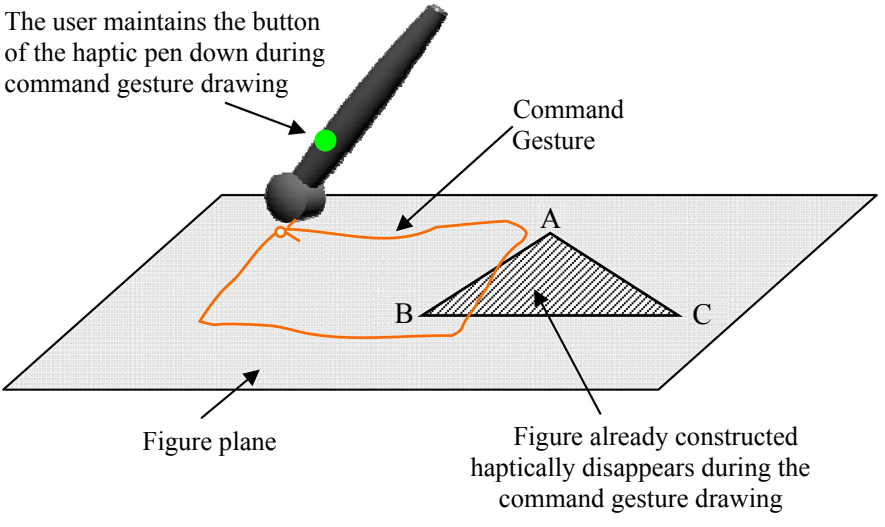


Fig. 2. Command gesture and figure exploration are performed in the same plane

[4] SYMBAL-latin recognition engine developed by Eric Petit and Sylvie Vidal (France Telecom R&D laboratory, Grenoble, France), partially based on theoretical work from LIP6, Paris 6 University (Marukatat, Artières & Gallinari, 2004), in a collaborative work with France Telecom R&D.

and draws a gesture, then it is interpreted as a command gesture. Gesture recognition is launched as soon as the user releases the stylus switch. And when the stylus switch is not held down, then user is only in haptic feeling mode.

With this solution, contextual command gestures can be performed, using spatial information at the beginning of the command gesture drawing. If the user starts his gesture near an existing element, this one is taken into account in the geometric figure construction process (if it is relevant for the command requested).

For example, for the construction of a circle, we need two elements: the centre and a point on the circle. If the beginning of the circle command gesture drawing is located nearby an existing point, this one is interpreted as the centre.

But gains in contextual information are balanced by losses in command-perception distinction. Indeed our new interaction technique amalgamates command (modifying an haptic object) and perception of an object (touching it only). Will the switch be enough to make the difference between perceptive only elements and commands elements?

Another point considering command gesture is the learnability of such gestural elements. As seen in 2), many command gestures systems exist, but their main drawback is that they are built on gesture visual representation: the draw of the gesture, not the gesture itself.

4 Experiments

Whether such a system can be used by visually impaired pupils is not a trivial question. First, we have to ensure that visually impaired pupils are able to learn how to achieve accurate command gestures using the haptic pen. Second, we want to validate such an interaction mode (based on a single haptic device) by asking pupils to solve a geometrical construction exercise. Those issues led us to carry out two different experiments described below[5].

4.1 First Experiment: Gestures Learning

Is the haptic modality through a pen relevant to teach gestures to visually impaired children? What kind of gestures can be easily learned and reproduced by the widest population of visually impaired pupils? Is our 2-dimensional gestures recognition engine, which proves very efficient for cursive letters writing, sufficiently adaptive to the intrinsic variability of our population?

4.1.1 Description

In order to test the ability of subject to reproduce with the haptic pen a gesture learnt with this pen, a set of eleven gestures has been defined according to some gesture complexity criteria previously identified in [6]: straight/curve, orientation, opened/closed. The idea is to cover a sufficiently heterogeneous set of gestures so that our results do not depend on one gesture family.

[5] This work is carried out in collaboration with the EREADV (*Etablissement Régional d'Enseignement Adapté pour Déficients de la Vue*) of Villeurbanne, a specialized school for the visually impaired near Lyon in France.

Each subject has to follow a scenario for each of the eleven predefined gestures:

- during the learning step, the haptic pen guides the hand of the subject along the gesture[6], that is repeated five times. The learning procedure is completed with oral indications from the system, and pupils are asked to tell what kind of shape they are going through (as soon as they recognize it).
- at the end of the five repetitions of the same gesture, the subject is asked to reproduce the learned gesture three times. Reproduction time is not constrained.

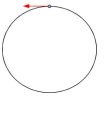
The presentation order of the eleven gestures is randomly determined for each subject. 19 people attend this first experiment: 8 blindfolded adults, 5 pupils suffering from amblyopia, 1 pupil blind from birth and 5 pupils who became blind.

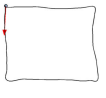
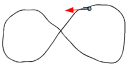
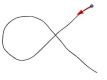
4.1.2 Results

The task is understood by every subject. All subjects succeed in perceiving and reproducing gestures, though some gestures appear to be harder to reproduce than others (as shown in Table 1).

This table shows the average score given by our recognizing engine on each gesture: score is given for each reproduction between 0 and 100, 0 means that the gesture has not been recognized, and 100 means that the reproduced gesture is exactly the same as the gesture model. It points out a gesture complexity notion which could lead to carry out some new experiments to systematically isolate gesture complexity criteria.

Table 1. Gestures classification from the best reproduced gesture to the worst

Name	6	11	5	7	9
Gesture					
Avg recog score	95,2	86,9	82,4	77,6	76,5
StdDev	15,5	24,3	15,9	33,9	32,7

8	4	1	2	3	10
					
73,3	67,0	65,2	63,9	56,5	51,9
36,8	36,2	28,9	34,2	33,1	39,2

[6] This method is inspired from a software called Télémaque ([13]) which aims at helping pupils learn handwriting.

As expected, these results confirm the importance of the semantic association: a gesture is less well learnt and reproduced if it has no meaning for the subject (gestures 2, 3 and 10). This effect is significant for every group considered (blindfolded, amblyopic and blind pupils).

Moreover this classification highlights the difficulty to reproduce some gestures already included in our interface (for example gesture 2 used to create a point).

4.2 Second Experiment: Performing a Typical Exercise

The second experiment asks the following question: is a pupil able to do a geometry drill (that corresponds to his geometry knowledge) using command gestures? We already showed that it is easily achievable by blind pupils using haptic and audio perception and keyboard commands [3]. This second experiment aims at showing that the same kind of activity can be successfully achieved through a gesture command interface and without any keyboard or menu.

4.2.1 Description

First the geometry drill and available command gestures (four gestures in total) are presented to the subject. A three-step exercise is proposed, instructions are given step by step. First, the pupil is asked to build a triangle: it is achieved by drawing a triangle.

Second, he has to build the midpoint of the segment [AC]. If he succeeds in making the gesture 2 (alpha) from any point of the segment [AC] (context sensitive action), the midpoint is automatically constructed and called I. At Last, he has to build the perpendicular bisector of [AC]. To do so, he has to draw the gesture 9 (context sensitive action also) and then has to click on the point I to select it by pressing the stylus switch button.

24 people attend this second experiment: 8 blindfolded adults, 9 pupils suffering from amblyopia, 2 pupils blind from birth and 5 pupils who became blind.

4.2.2 Results

The global results are very encouraging because all subjects succeed in finishing the proposed geometrical construction. This experiment validates the use of a full haptic pen based interface to make geometry exercises. Nevertheless, three visually impaired pupils have been helped to accomplish at least one of the steps. Gesture alpha proved to be very difficult, particularly for pupils suffering from amblyopia and late blind pupils (an average of almost 4 tries for each of them). After examination, this difficulty is not due to the contextual aspect of the gesture (unlike the gesture 9 for which the consideration of the context mislead some subjects).

4.3 Discussion

Those two experiments prove to be very fruitful, the first one confirms that our system can work as an autonomous system because the haptic-based gestures learning is now validated as an efficient way to learn command gestures; and the second one validates our approach: visually impaired pupils succeed in a geometry figure construction task using one single device (the haptic pen) for action and perception.

Moreover, they have highlighted some points about our choices of gestures. For example, we foresaw the use of the alpha gesture to create a point. Both experiments show it is not a good choice: we finally decide to associate a simple click on the stylus switch to this command.

Lastly, a detailed examination of the reasons why errors occurred lead us to a better understanding of the specificity of blind pupils: perception of gestures using the only sense of touch is very different from visual representation. For example gestures 11 and 8 are close for a blind pupil (3 strokes with similar directions) whereas they are visually quite different.

5 Current Work

Thanks to the success of these experiments, one version of SALOME based on command gestures has been installed at the EREADV, the specialized school for visually impaired children, in December 2006. Now any teacher can use it with a pupil within his geometry lesson. Unfortunately the gesture learning module (pupil's hand being guided by haptic pen) has not been integrated yet in this version. First feedbacks (3 pupils) are very encouraging as they succeed in using interface, gestures being learnt with teacher's help. Though several drawbacks have already been pointed out:

- learning gestures step is long and hard according to pupils: this confirms the need to integrate the learning gestures module.
- some gestures appear to be not appropriate. For example, two pupils did not manage to perform the erase command. This command is launched by the gesture "e", which is a typically graphical representation of a sighted person, but that has no semantic meaning for a blind pupil and that is gesturally similar to the circle gesture (whereas their graphical representation is quite different). This confirms the huge difference between graphic similarity and gestural similarity, and thus the need to define the set of command gestures on gestural criteria only.

6 Future Work

Our work placing the handicapped people at the centre of the design process allows our interface to be usable by everybody ("design for all"). Moreover Gentaz and al. [1], [2] showed that the haptic exploration of letters increases the children reading and handwriting acquisition. We can suppose similar effects in geometry learning. Thus we plan to broaden our experiments to sighted pupils and test the effect of our haptic-based interface on geometry learning in standard education. The underlying idea is that providing concrete experiences with intangible concepts and ideas can help pupils to better comprehend these concepts and ideas.

In addition, thanks to a research collaboration with a company that develops and distributes an interactive geometry software for sighted, we plan to integrate their SDK in SALOME in order to benefit from their teaching method and dynamics (mainly the possibility to modify the figure, all geometric properties being preserved).

7 Conclusion

Our past and present work on SALOME leads to a prototype which is: (1) efficient, pupils can perform geometry exercises with a few simple commands; (2) user-friendly, no more keyboard has to be used to perform a construction and a single plane to both build and feel a figure makes the interaction easier; (3) autonomous, gestures can be taught through an haptic pen which is a particularly relevant method for blind pupils.

It is essential for us that SALOME is currently tested and used by the impaired pupils who lead us to design and develop this software. It gives the opportunity to every impaired pupil to work on a computer tutorial as any others pupils. The advantages are the flexibility of the tool compared to the usual teaching method (pen and plastic sheet), and the possibility to collaborate with a sighted pupil by using the same interface designed to be accessible to everybody. All this promotes visual impaired pupils' inclusion.

References

1. Bara, F., Gentaz, E., Colé, P., Spenger-Charolles, L.: The visuo-haptic and haptic exploration of letters increases the kindergarten-children's understanding of the alphabetical principle. Cognitive Development 19, 433–449 (2004)
2. Palluel-Germain, R., Bara, F., de Boisferon, H.A., Hennion, B., Gouagout, P., Gentaz, E.: A visuo-haptic device - Telemaque – increases the kindergarten children's handwriting acquisition. IEEE WorldHaptics (in press)
3. Zijp-Rouzier, S., Petit, E.: Teaching geometry to visually impaired people using haptics and sound. Human-Computer Interaction (2005)
4. Minogue, J., Jones, M.G: Haptics in education: exploring an untapped sensory modality. Review of Educational Research. Education Periodicals 76(3), 317 (2006)
5. Rouzier, S., Hennion, B., Perez-Segovia, T., Chêne, D.: Touching geometry for visually impaired pupils. EuroHaptics Proceedings, Munich, Germany, pp. 104–109 (2004)
6. Klok, M., Uzan, G., Chêne, D., Zijp-Rouzier, S.: Blind gestural interaction: an experimental approach. In: Gesture Workshop Proceedings, Valoria, Université de Bretagne Sud, Vannes, France (2005)
7. Gillespie, R.B., O'Modhrain, M.S., Tang, P., Zaretzky, D., Pham, C.: The virtual teacher. In: Proceedings of the ASME IMECE, vol. 64, p. 171–174 (1998)
8. Mertz, C.P., Lecoanet, P.: GRIGRI: Gesture Recognition on Interactive Graphical Radar Image. Progress in Gestural Interaction, Springer. In: Proceedings of Gesture Workshop '96, pp. 187–198, University of York, UK (1997)
9. Kamp, J.-F., Poirier, F., Doignon, Ph.: Control of In-vehicle Systems by Gestures. Gesture Based Communication in Human-Machine Interaction. In: Braffort, A., Gibet, S., Teil, D., Gherbi, R., Richardson, J. (eds.) GW 1999. LNCS (LNAI), vol. 1739, pp. 159–162. Springer, Heidelberg (2000)
10. Mertz, C.P., Vinot, J-L., Etienne, D.: Entre manipulation directe et reconnaissance de l'écriture: les gestes écologiques. ErgoIHM (Octobre 2000) Biarritz, France (2000)
11. Mouse Gestures (Mozilla Firefox) (050105) http://optimoz.mozdev.org/gestures/installation.html
12. Mouse Gestures (Opera) (050105) http://www.opera.com/features/mouse/
13. Hennion, B., Gentaz, E., Gouagout, P., Bara, F.: Telemaque, a new visuo-haptic interface for remediation of dysgraphic children. In: Proceedings of the first WorldHaptics Conference, Pisa, Italy, Mars (2005)

Unified Design of Universally Accessible Games

Dimitris Grammenos[1], Anthony Savidis[1], and Constantine Stephanidis[1,2]

[1] Foundation for Research and Technology - Hellas (FORTH), Institute of Computer Science, Human-Computer Interaction Laboratory, Heraklion, Crete, GR - 70013 Greece
[2] University of Crete, Department of Computer Science
{gramenos,as,cs}@ics.forth.gr

Abstract. This paper describes how the Unified Design method, originally conceived as a means for developing universally accessible user interfaces, can be adapted and applied to the development of universally accessible games. The basic steps for applying the method are presented and explained through illustrative examples. Furthermore, the key differentiations between designing turn-based strategy games and action games are highlighted and the related impact to the application of the design method is explained.

Keywords: Universally Accessible Games, interaction design methods, Design for All, polymorphism.

1 Introduction

Until now, little attention has been paid to the development of computer games that can be potentially played by all gamers, independently of their individual characteristics, requirements, preferences and abilities. In particular, there are no computer games that can be concurrently played among able and disabled people, either remotely or sharing the same computer, with the minor exception of a few games that can be played both by visually impaired and sighted players, like All inPlay card games[1] and the 3D shooter Terraformers[2].

This paper introduces a novel approach to creating games which are proactively designed to be concurrently accessible by people with a wide range of diverse requirements and / or disabilities. In this context, a structured design method, based on Unified User Interface Design [1], is suggested. The method follows a step-by-step, top-down approach, starting with a high-level abstract task definition process and eventually leading to the creation of a complex - but well-structured - design space, populated with numerous interweaved physical designs.

2 Background and Related Work

From a technical point of view, two main approaches have been adopted to address the issue of computer game accessibility:

1. Inaccessible games become operationally accessible through the use of third-party assistive technologies, such as screen readers, mouse emulators, or special input

[1] http://allinplay.com/games.html
[2] http://www.terraformers.nu/index.php?option=com_content&task=view&id=34&Itemid=28

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 607–616, 2007.
© Springer-Verlag Berlin Heidelberg 2007

devices. In practice there are serious barriers and bottlenecks inherent in the absence of compatibility efforts during the development of computer games and assistive technology systems. However, even when some sort of compatibility is achieved, this is typically the result of either customized low-level adaptations (i.e., hacking) or pure coincidence, rather than the outcome of appropriate design considerations.

2. Accessible games are developed from scratch, however, targeted merely to people with a particular disability, such as audio-based games for blind people[3], and single-switch games for people with severe motor impairments on the upper limbs[4].

Following the first approach we typically accomplish a very limited form of accessibility, as well as poor interaction quality and usability. Through the second approach, being the most promising, we have to cope with two key drawbacks: (a) there is a significant tradeoff between the cost of developing high quality accessible games and the expected return on investment, assuming the target user group reflects a limited market population; and (b) there is an apparent hazard due to the potential segregation between able and disabled gamers, essentially leading to social exclusion.

In order to overcome the limitations of existing approaches towards game accessibility, the Human–Computer Interaction Laboratory of ICS-FORTH, following the principles of Design for All [2], has introduced the concept of Universally Accessible Games [3] (or, UA-Games), as an effective technical approach to achieve game accessibility coupled with high interaction quality, also putting forward the objective of creating games that are concurrently accessible to people with diverse abilities.

At present, in the context of the UA-Games research activity[5] of the HCI Lab of ICS–FORTH, two games have been developed: *UA–Chess* [3], a universally accessible web–based chess; and *Access Invaders* [4], a universally accessible multiplayer / multiplatform version of Space Invaders.

3 The Design Process

A key prerequisite to effectively accommodate the particularly broad spectrum of diverse interaction requirements imposed by universally accessible games is to firstly design the interactive game space at an abstract level, in a representation-independent way, eliminating all references to the physical-level of interaction (e.g. input / output devices, rendering, low-level dialogue). Once this is accomplished, the next step is to appropriately capture the lower-level design details, incrementally specializing towards the physical level of interaction by addressing particular user characteristics. To this end, a design approach capable to represent an open set of alternative physical designs under a common abstract design umbrella is the Unified User Interface Design method [1]. This method reflects a process-oriented discipline emphasizing abstract task definition with incremental *polymorphic* physical specialization.

[3] e.g., http://audiogames.net
[4] e.g., http://oneswitch.org.uk/4/games/reviews/index1.htm
[5] http://www.ics.forth.gr/hci/ua-games

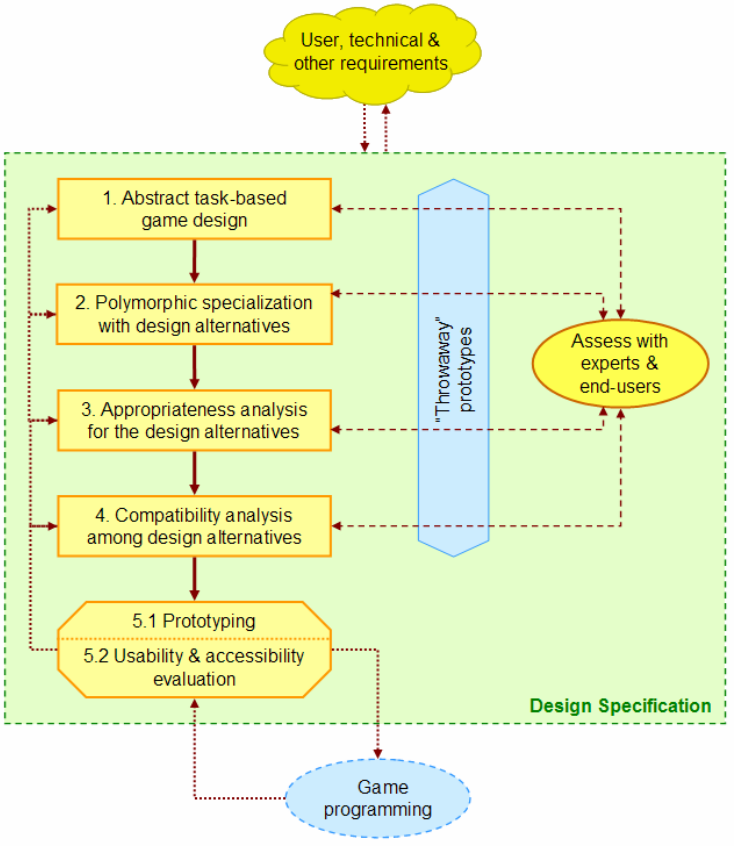


Fig. 1. Applying Unified Design to the development of universally accessible games

Since Unified Design was originally targeted to the creation of accessible user interfaces, it had to be adapted to cater for the intrinsic characteristics and particular needs of game design. The basic steps in applying Unified Design to the development of accessible games are summarised in Fig. 1. As shown, it is a highly participatory, user-centred, iterative process, as: (a) throughout the overall lifecycle, the direct involvement of several representative end-users (gamers) with diverse characteristics, as well as domain experts (usability, accessibility, gaming, etc.) is promoted for the continuous assessment of the design outcomes in each step; and (b) it is possible to return to a previous design step, in case, for instance, more information is required, some design artifacts have to be revisited, or the design parameters must be further specialized.

Quite often, in order to evaluate the decisions made at a specific step, or to weigh alternatives before committing to them, it is required to quickly create small-scale temporary prototypes, known as “throwaway” prototypes. These may range from rough hand-made sketches to simple programs. Prototyping [5] is an essential part of the iterative design process since it provides a low cost, tangible means for gathering

early and meaningful user feedback, and, at a later stage, can also serve as a common reference point, as well as a concrete, unambiguous, documentation medium for communicating design specifications to game programmers.

At this point, it should be noted that game programmers are also involved in the whole process with a two-fold role: (a) they provide input about technical requirements and restrictions, as well as about the feasibility and cost of alternative design solutions; and (b) they develop and "tweak" the required electronic prototypes.

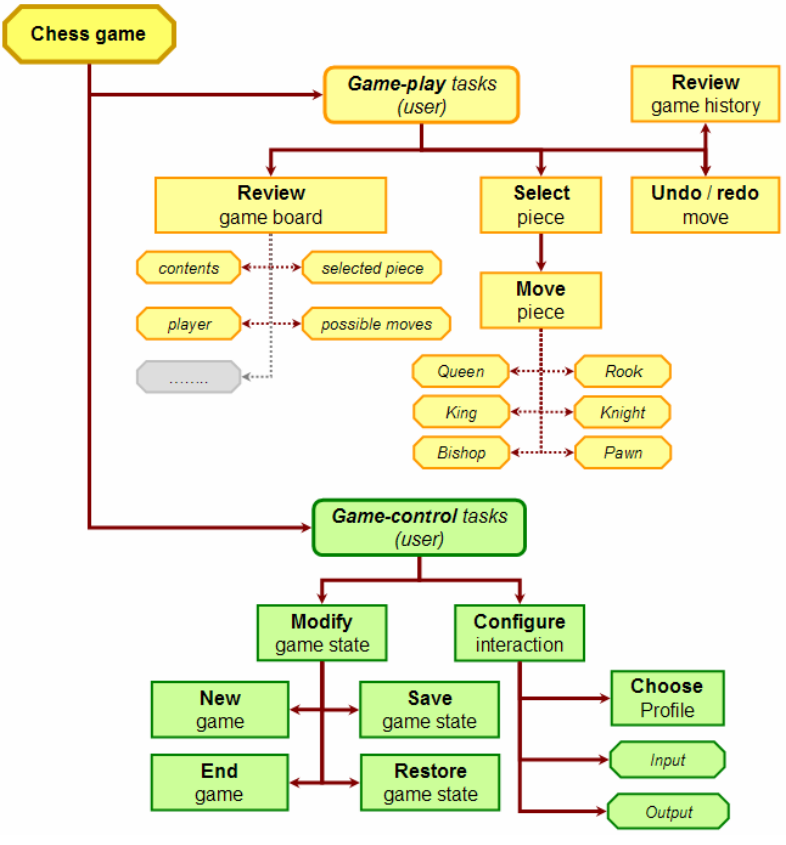


Fig. 2. Abstract task decomposition for the game of chess

3.1 Abstract Task-Based Game Design

The goal of this first step is to breakdown the high-level tasks performed by people when playing the particular game – irrespectively of the medium they use to play it – as well as the things they do, the things they act on and the things they need to know. The hierarchical decomposition adopts the original properties of hierarchical task analysis [6] for incremental decomposition of user tasks to lower level actions. In this

context, it is essential to focus on the basic logical game activities and constituents identifying their semantic attributes and relevant regulations independently of the way these can be physically instantiated to be accessible or usable to particular user groups.

As an indicative example, the result of the decomposition for the game of chess is illustrated in Fig. 2, where tasks are divided in two broad categories: (1) *game-play tasks*, comprising user actions directly related to the game goal and content (i.e., the board and the pieces); and (2) *game-control tasks*, which include "peripheral" user actions that affect the game state and the way that player-game interaction is performed.

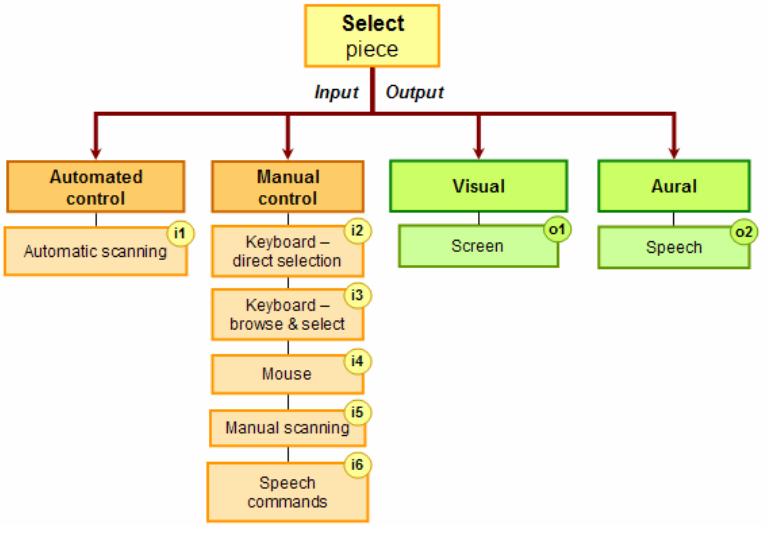


Fig. 3. Example of mapping abstract tasks to alternative interactive designs

3.2 Polymorphic Specialization with Design Alternatives

In this step, the abstract tasks resulting from the previous one are mapped to multiple low-level, physical alternative interactive designs, meeting target user attributes. In this context, accessibility barriers that can possibly emerge in each task when performed by a particular user group are identified and suitable alternative interaction methods and modalities are selected. An example of how the abstract task entitled "Select piece" (from the example illustrated in Fig. 2) can be mapped to alternative low-level, physical, alternative interactive designs is presented in Fig. 3.

3.3 Appropriateness Analysis for the Design Alternatives

A matrix is constructed correlating the perceived appropriateness of each selected design alternative for every user attribute. The rows of the matrix represent distinct user attributes while its columns alternative interactive designs. Each cell, depending on the suitability of the particular design for the specific user attribute, is filled with

one of the symbols depicted in Table 1. The alternative interactive designs appropriateness matrix can be filled in by reviewing related literature, using previous know-how in the field, as well as by questioning domain experts and representatives of the target user groups.

Table 1. Alternative interactive designs appropriateness symbols

Symbol	Meaning
✓ (ideal)	Explicitly designed for this user attribute.
⬆ (appropriate)	Suitable, but possibly not the best choice.
⬇ (could be used)	If nothing else is available, it can be used, though not recommended.
✕ (inappropriate)	Totally inappropriate, will result in posing an accessibility barrier.
= (neutral)	Does not have any effect on the particular user attribute.

	Design alternatives							
User attributes	i1	i2	i3	i4	i5	i6	o1	o2
Full vision	=	=	=	=	=	=	✓	⬆
Low vision	⬇	⬆	⬆	⬇	⬇	✓	⬇	✓
Blind	⬇	⬆	✓	✕	⬇	✓	✕	✓
No hand-motor impairment	⬇	✓	⬆	✓	✕	✓	=	=
Uses multiple switches	⬇	✕	⬇	✕	✓	✓	=	⬇
Uses single switch	✓	✕	✕	✕	✕	✓	=	⬇
Mild cognitive impairment	⬇	⬇	⬇	⬆	⬇	⬇	⬆	⬆
Expert	⬇	✓	⬇	⬆	⬇	✓	=	⬇
Novice	⬇	⬇	⬆	✓	⬇	⬆	✓	✓

✓: ideal ⬆: appropriate ⬇: could be used ✕: inappropriate =: neutral

Fig. 4. Example of appropriateness matrix for the design alternatives

A basic design goal is that for every abstract task, there should be at least one "ideal" or "appropriate" input and one output design alternative for each user attribute and target *user profile*. A *user profile* is a collection of user attributes (e.g., novice, sighted, hand-motor impaired gamer). The appropriateness of a design alternative for a specific user profile can be inferred by merging the corresponding rows of the matrix that contain attributes of this profile, as follows:

- If the design alternative is *inappropriate* for any of the user attributes, then it is deemed as inappropriate for the entire profile.
- If the design alternative is *neutral* for a specific attribute, it means that it is not related to this attribute and thus, does not affect the design's appropriateness for it.
- In all the other cases (i.e., *ideal*, *appropriate*, *could be used*), the lowest appropriateness value supersedes all the others.

Thus, for example, Fig. 5 illustrates the appropriateness matrix for the case of a low vision, novice player that can use just a single switch. As can be seen, this is a particularly difficult case, since the available solutions are very limited and not optimal. Nevertheless, it can still be ensured that the game is accessible in this case.

	Design alternatives							
User attributes	i1	i2	i3	i4	i5	i6	o1	o2
Low vision, novice, single-switch	↓	×	×	×	×	↑	↓	↓

Fig. 5. Appropriateness matrix for a low vision, novice player that can use just a single switch

	i1	i2	i3	i4	i5	i6
i1		✓	×	✓	×	✓
i2	✓		✓	✓	✓	✓
i3	×	✓		✓	×	✓
i4	✓	✓	✓		✓	✓
i5	×	✓	×	✓		✓
i6	✓	✓	✓	✓	✓	

✓: compatible ×: incompatible

Fig. 6. Compatibility matrix example for the alternative input designs of Fig. 3

3.4 Compatibility Analysis Among Design Alternatives

When alternative interactive designs have been identified, it is essential that cases where two, or more, alternatives are mutually incompatible are pinpointed, so that they can be avoided. To this purpose several related *compatibility matrices* need to be devised. A *compatibility matrix* has as rows and columns all the alternative interactive designs that can potentially be concurrently active at a particular point in time. If two designs are compatible then the corresponding cell is filled with a green tick (✓), else with a red X (×). Fig.6 presents a compatibility matrix created for the alternative input designs of the example presented in Fig.3.

3.5 Prototyping, Usability and Accessibility Evaluation

As soon as a "stable" version of the whole (or just a discrete part of) game design is available, indicative electronic prototypes of the game can be developed showcasing the alternative interactive properties of its user interface for the different target user groups. The usability and accessibility of these prototypes should definitely be

evaluated with representative end-users. In this respect, a quick, handy and very effective informal evaluation method is *thinking aloud* [7]. According to this method, an evaluator (or sometimes even more) observes a gamer interacting with the system, asking vocalisation of thoughts, opinions and feelings, in order to understand the gamer' s mental model, the way she thinks and performs tasks and find out any mismatches between the user's mental model and the system model. Conversations are usually recorded so that they can be analysed later on. Furthermore, to support the evaluation process, a list of indicative tasks is used, that prompts participants to explore the full game functionality available. After playing the game, a small debriefing session can be held, where participants are asked about their overall impression of their interaction with game and personal preferences, likes or dislikes, as well as for suggestions regarding potential improvement and modifications.

The outcomes of this step, can considerably aid in validating, correcting and updating design decisions, as well as in developing new ideas for improving the accessibility and playability of the final game. When the design specification is considered as "mature" it can then be propagated for further development. Of course, as parts of the game and its functionality are being developed it is highly desirable to regularly perform usability and accessibility testing.

4 Abstract Task-Based Game Design and Action Games

A basic design differentiation between a turn-based board game, such as chess that was used as an example up to now, and action games is the "degrees of freedom" along which the game can be modified in order to become accessible. More specifically, in the case of board games, only the user interface (i.e., the way the game is presented and controlled) can be adapted to better match a particular player's characteristics. The game's rules (logic) and content are fixed and any deviations are not possible. Then again, action games are more flexible in this dimension, since they usually have a main goal, but they do not impose restrictions on how this can be achieved and do not have any globally established strict rules and specific content.

To provide an illustrative example, when making a chess game accessible, the possible adaptations can only affect how the board and pieces are rendered and the way the player can select and move the pieces. The type or number of the pieces, the rules they follow for moving, or what a player has to do to win can not be changed (since this would be a different game). On the contrary, when creating an accessible version of Space Invaders, beyond changing how the player's spaceship is controlled and presented, it is possible to completely revamp the characteristics of the attacking alien ships (e.g., number, speed, firepower, size) and even the rules of the game (e.g., allow the player to destroy any alien, but only a specific alien to destroy the player, change the initial number of the player's "lives").

Thus, a key difference in relation to the aforementioned design example of the chess game is that, at the stage of identifying accessibility barriers related to the game's interface, barriers that stem from the game's content and rules should also be identified, as well as possible design strategies for overcoming them. In this respect,

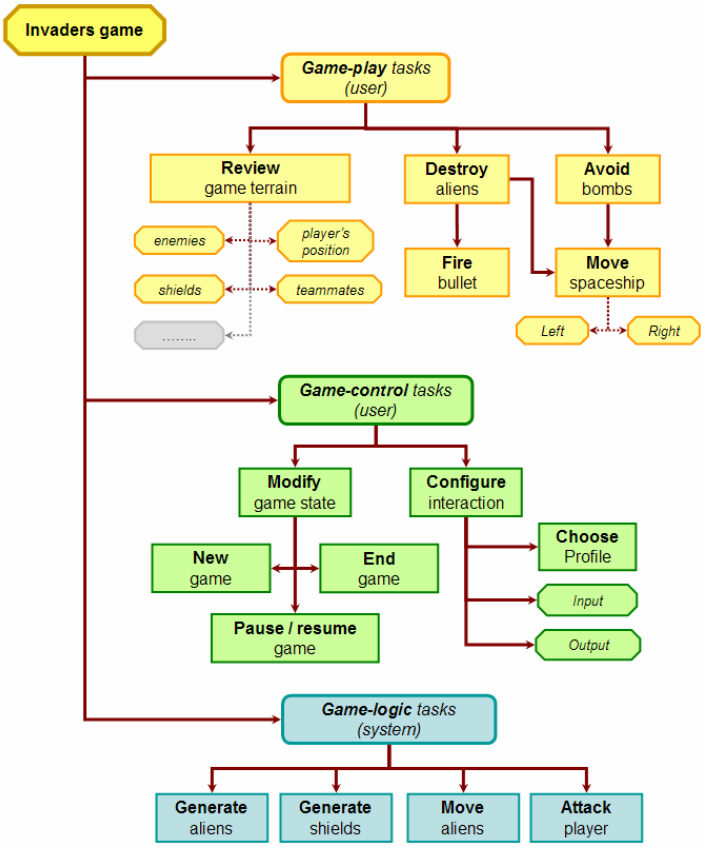


Fig. 7. Abstract task decomposition of a "Space Invaders" type of game

the abstract task decomposition of a "Space Invaders" type of game is illustrated in Fig. 7. This time, tasks are divided in three – instead of two – categories:

1. *game-play tasks*, comprising user actions directly related to the game content;
2. *game-control tasks*, which include "peripheral" user actions that affect the game state and the way that player-game interaction is performed;
3. *game-logic tasks*, which are performed by the system in order to create and control the various active game elements.

5 Conclusions

The design of universally accessible games is an inherently demanding and challenging task, requiring the management of a very large design space, which, even when the game design is mostly completed, may still grow dynamically as new diverse user groups or varying environments of use are accommodated in the design

process. Additionally, it entails the mapping and transformation of design parameters to coherent, highly usable and accessible interaction designs.

To this end, analytical design methods, such as the one presented in this paper, are anticipated to become necessary tools for capturing, representing and managing the large and complex design spaces of universally accessible games, addressing the compelling requirements for customization, accessibility and high quality of interaction.

The presented design method has been formulated, applied, tested, refined, and documented in its present form through the development of two computer games in the context of the UA-Games research activity of ICS–FORTH (see section 2 "Background and Related Work").

References

1. Savidis, A., Stephanidis, C.: Unified User Interface Design: Designing Universally Accessible Interactions. International Journal of Interacting with Computers 16(2), 243–270 (2004)
2. Stephanidis, C., Salvendy, G., Akoumianakis, D., Bevan, N., Brewer, J., Emiliani, P.L., Galetsas, A., Haataja, S., Iakovidis, I., Jacko, J., Jenkins, P., Karshmer, A., Korn, P., Marcus, A., Murphy, H., Stary, C., Vanderheiden, G., Weber, G., Ziegler, J.: Toward an Information Society for All: An International R&D Agenda. International Journal of Human-Computer Interaction 10(2), 107–134 (1998)
3. Grammenos, D., Savidis, A., Stephanidis, C.: UA-Chess: A Universally Accessible Board Game. In: Salvendy, G. (ed.) Proceedings of the 3rd International Conference on Universal Access in Human-Computer Interaction, Las Vegas, Nevada, USA, Lawrence Erlbaum, Mahwah (2005)
4. Grammenos, D., Savidis, A., Georgalis, Y., Stephanidis, C.: Access Invaders: Developing a Universally Accessible Action Game. In: Miesenberger, K., Klaus, J., Zagler, W., Karshmer, A.I. (eds.) ICCHP 2006. LNCS, vol. 4061, pp. 388–395. Springer, Heidelberg (2006)
5. Floyd, C.: A Systematic Look at Prototyping. In: Budde, R., Kuhlenkamp, K., Mathiassen, L., Zullighoven, H. (eds.) Approaches to Prototyping, Springer, Heidelberg (1984)
6. Johnson, P., Johnson, H., Waddington, P., Shouls, A.: Task-related knowledge structures: analysis, modeling, and applications. In: Jones, D.M., Winder, R. (eds.) People and Computers IV,, pp. 35–62. Cambridge University Press, Cambridge (1988)
7. Nielsen, J.: Usability Engineering. Academic Press, Boston, MA (1993)

Universal Design Within the Context of e-Learning

Andrina Granić and Maja Ćukušić

Faculty of Natural Sciences, Mathematics and Education, University of Split
Nikole Tesle 12, 21000 Split, Croatia
andrina.granic@pmfst.hr, maja.cukusic@st.t-com.hr

Abstract. When designing an e-learning system it is important to consider not only technical and resource related aspects, but user characteristics along with pedagogical and contextual issues as well. This paper proposes an approach to universal design within the e-learning environment able to address issues related to learner-centred design paradigm, context of use approach, individualized approach, pedagogical framework and guideline framework. A refinement from methodological point of view enhanced with the exemplification of approach's appliance is required.

Keywords: universal design, e-learning, context of use, learner-centred design, pedagogical framework.

1 Introduction

An evolution of computer from computational machine to "knowledge machine" influences a development of variety of new computer-mediated human activities, among which e-learning merits special attention. Placed on the crossroad between information and communication technology (ICT) and education, e-learning is considerably changing traditional definition of an educational system. However, despite so much publicity and activity, the progress in the field has been very slow. Related problems are mostly associated with poor design of e-learning applications [34] as well as still inadequate research base for e-learning evaluation [1]. Furthermore, although computers are being used at different levels of teaching process (as the subject of teaching as well as a tool for supporting the teaching process) and despite decades of research, their use for tutoring (as the teacher itself) in everyday teaching environment has been quite limited [16].

On the other hand, the objective of design for all or universal design is to provide accessible and usable information society technology (IST) products and services, raising so called *good user-centred design* paradigm to a more encompassing concept of addressing the needs of all potential users, *cf.* [31]. Therefore, an employment of universal design within the e-learning context promotes individualization and end-user acceptability, ensuring that usability and accessibility should be a design concern, thus avoiding the need for afterwards adaptations. Unfortunately, studies have regularly shown that the accessibility of web-based applications in general falls short of an acceptable level [25] and most of existing efforts related to accessible e-learning ones propose guidelines that primarily address technical accessibility issues [6].

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 617–626, 2007.
© Springer-Verlag Berlin Heidelberg 2007

Additionally, when considering e-learning applications it has been claimed that usability assessment needs further consideration of learning perspective [20], [28], although some authors propose applying heuristics without further adjustment to the e-learning context, e.g. [7], [21]. Nevertheless, in e-learning we do not need user interfaces that support "doing tasks", but interfaces that support "learning while doing tasks", *cf.* [12]. Obviously, there should be a synergy between the learning process and a user's/learner's interaction with the application [29], hence taking into account the different ways user's learn and ensuring their natural and flexible interactions as well.

Our research, focused on e-learning systems which emulate human teacher in the process of learning and teaching, brought about comparable assumptions. Although employed in the teaching process of a few real subject matter, they have been continually evaluated concerning their efficiency in the learning process and usability of their interface design. Related studies and empirical assessments however still mainly involve usability aspects, *cf.* [10], yet not addressing issues concerning accessibility and learning perspective as well. Applied methodology, based on the combination of behaviour and opinion founded measures, linked scenario-based user testing with a kind of "less formal" heuristic evaluation, *i.e.* guideline evaluation. A major strength in such an approach is the chance to supplement results from both the heuristic evaluation and the empirical user-based one, enhanced by users' feedback on their comfort while working with the system. Nevertheless, the design of accessible and easy to use system able to address the needs of all potential users requires additional considerations in the context of e-learning environments.

The rest of the paper is structured as follows. Section 2 provides an overview of a number of pertinent approaches to efficient and usable e-learning design. These studies, along with our research and related findings, denote irreplaceable input for the description of the proposed "frame" for an inclusive design within the e-learning context. The approach along with its foremost aspects is elaborated in Section 3. Future research directions related to its refinement and application are discussed in Section 4. The same section concludes the paper.

2 Related Work

With the aim of design e-learning systems with intuitive interfaces, ones which do not get in the way to effective learning and teaching, recent HCI research is starting to address a learning perspective as well. Brief insight into few studies along with identified main aspects describing best practice for interface design of e-learning systems is presented in the following.

A holistic approach proposed by Phipps and Kelly [22] focuses on accessibility aspect and is comprised of few aspects: usability issues, making learning accessible, making courses accessible as well as adapting to individual, local political and cultural factors. The framework recognizes that accessibility of the system is primarily about people and not about technology. Hence, the design should take into account (i) accessibility, (ii) usability, (iii) learning outcomes, as pedagogic issues and aims of the e-learning resources, (iv) learner needs, (v) local factors, (vi) infrastructural and resources issues and finally (vii) quality assurance. Viewpoint outlined by De Marsico

et al. [6] puts more emphasis on a joint design activity of e-learning system users and multidisciplinary experts who possess different skills related to didactical, pedagogical, technical and usability aspects, considering users with disabilities as well. In such context the roles and tasks of each participant in every successive phase of the e-learning development lifecycle should be determined. Good and Robertson [9] argue that current learner-centred design frameworks do not consider the full range of issues necessary for successful design. Their learner-centred design framework named CARSS, specifically designed for child learners, embraces context, activities, roles, stakeholders and skills. Based on literature review, Zaharias [34] summarizes the most commonly identified research dimensions regarding quality and e-learning design issues, pointing out efforts in:

- implementation of learner-centred design paradigms,
- implementation of effective pedagogy for the design of e-learning courses and subsequent development of instructional design guidelines, along with
- guidelines and frameworks for quality assurance and evaluation.

Apparently common aspects appear in each study. Nevertheless, each approach to the learning and teaching system design emphasises one specific aspect in particular. Thoroughly considering important factors proposed by multidisciplinary design experts and bearing in mind findings from the related studies, we have devised approach to inclusive design within the context of e-learning.

3 An Approach to Inclusive Design in e-Learning Environment

The frequently low level of usability and accessibility of e-learning systems indicates that focusing on the adoption of usability and accessibility guidelines is not sufficient for a truly design for all. As a result, simple appliance of usability heuristics, for example Shneiderman's golden rules or Nielsen's heuristics, cannot effectively be used because they fail to address the specific challenges of learner-centred design and the issue of integration of usability and learning [28]. Additionally, simple adoption of W3C guidelines in order to assure accessibility is just not enough, because the vast majority of the available accessibility guidelines are formulated either as general design principles or low level specific recommendations, not focusing on the needs of the users [33]. Evidently there is a need for further research and empirical assessment since an established set of heuristics [35] as well as a joint evaluation methodology for e-learning applications do not yet exist [2], [5].

Consequently, the application of design for all in an educational environment elevates end-user acceptability and customization of learning to individual needs, ensuring that usability and accessibility should be a concern of design. For that reason the design of accessible and easy to use e-learning system, able to address the needs of all potential users despite usability and learning perspective, requires additional considerations. The main issues regarding universal design related to e-learning systems have to embrace:

- *learner-centred design paradigm*; the same practices followed by HCI community (employment of user-centred iterative design and repeated evaluation) must be used to ensure learnability, major issue for e-learning, according to Don Norman;

- *context of use approach*; the main measure of an e-learning system is whether it goes with its context of use in terms of four components of HCI model [23]: whether the user for whom it is designed can use it with acceptable levels of usability and accessibility, for the tasks that s/he needs to do, in the local environment in which these tasks take place, using the available technologies;
- *individualized approach*; the consideration of the users' different individual characteristics relevant to learning styles and preferences, such as social and cultural background, technical experience, interests, goals, physical/cognitive abilities and a like, fosters individualization and end-user acceptability;
- *pedagogical framework*; covering at least three aspects: pedagogies which underpin the subject matter in e-learning environment, content and information for learners; and ways in which the subject matter is delivered to learners (delivery strategies).
- *guideline framework*; the employment of usability and accessibility guidelines for e-learning quality assessment, ensuring that usability and accessibility should be a design concern, thus avoiding the need for later adaptations or specialized design.

Identified aspects of the proposed approach along with addressed issues concerning universal design within an e-learning context are illustrated in Figure 1 and elaborated hereinafter.

3.1 Learner-Centred Design Paradigm

The current research in HCI field acknowledges that understanding users' needs is the core of a successful design of IST products and services. This naturally leads to user-centred design approaches, a philosophy which places the users at the centre of design [19] and a process that focuses on cognitive factors (such as perception, memory, learning, problem-solving, *etc.*) as they come into play during users' interactions with application [34].

Moreover, in order to "make people more effective learners", *i.e.* to take into account the unique needs of users as learners, a shift from user-centred to learner-centred design is needed [27]. It can range from attempts to design with the needs of the learner at the forefront towards involving the learner at various stages of the design process [9]. Such approach entails understanding and considering who is the user, what are her/his needs, what we want her/him to learn, how is (s)he going to learn it and how are we going to support her/him in achieving the learning objectives. Accordingly, a variety of learners' types must be considered due to characteristics revealing user individual differences like personal learning styles and strategies, diverse experience in the learning domain, as well as previously acquired knowledge and abilities. Many authors have attempted to provide comprehensive lists of additional needs for specific educational domains, but [26] concisely summarize them under broad categories of universal applicability.

In brief, in order to address the distinctive needs of users as learners, learner-centred design attempts to provide guidelines to augment the user-centred design framework.

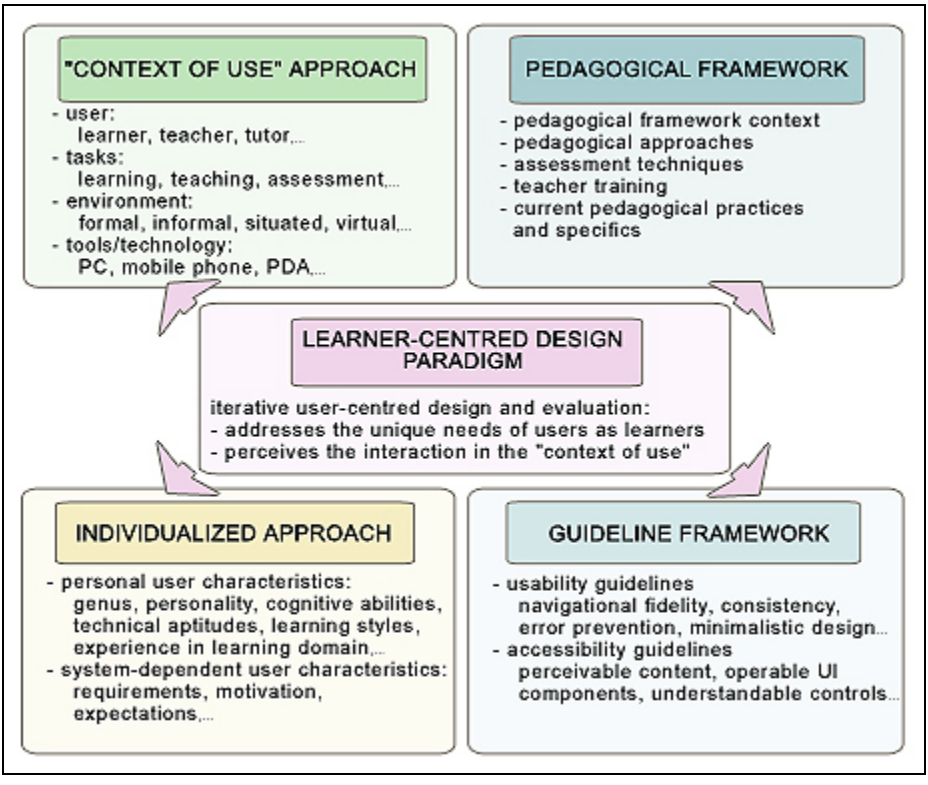


Fig. 1. An approach to universal design within the context of e-learning

3.2 "Context of Use" Approach

An understanding of the context of use forms a useful input in order to cope with the concept of design for all, an inclusive and proactive approach seeking to accommodate diversity in the users and usage context. The notion of context of use, defined by ISO 9241 as "… the nature of the users, tasks and physical and social environments in which a product is used" [13], is already well recognized in the concept of universal design e.g. [14], [30], and closely related concept of universal usability [24]. On the other hand, it has been claimed that despite the fact that the power of modern IST artefact have been increased by a factor of thousand over the past two decades, user interface design have mostly stayed the same [3]. Regardless of advances in HCI field, few innovations had get through into new information society systems for all. It seems that such problem can only be overcome by switching from a focus upon the design of user interfaces to the design of interactions themselves and to the "wider view of interaction in the context of use" [*ibid.*].

Consequently context of use, as an important concept of everyday life in general, can be seen as a description of the actual conditions under which information society technologies are under assessment or will be used in a usual day by day working situation. From the educational perspective in particular, the effective use of context

in an e-learning system goes in terms of four components of a HCI model [23]: whether (i) the user for whom it is designed can use it with acceptable levels of usability and accessibility, for (ii) the tasks that s/he needs to do, in (iii) the local educational, socio-cultural and organizational environments in which these tasks take place, using (iv) the available technologies (see Figure 1).

3.3 Individualized Approach

As already emphasized, present HCI research puts the user, the individual at the centre of all developments, stressing the importance to design IST products and services for human needs. Therefore, the role of an intuitive user interface and a flexible interaction suited to different needs, preferences and interests becomes even more important for the users' success, as users with a wide variety of background, skills, interests, expertise, goals and learning styles *cf.* [4], [8] are using computers for quite diverse purposes. For that reason no single interface, though designed to be both easily learnable and highly efficient, will satisfy every user.

Yet, the progress in the field of e-learning has been very slow and "… although technology is often touted as the great salvation of education – an easy way to customize learning to individual needs – it rarely lives up to this broad expectation" [11, p. 398]. It seems that too much of the research is being driven by technical possibilities, while paying inadequate attention to the area of application. The result is in an over-ambitious and pre-mature attempt to eliminate the teacher's role in educational environment [16]. Huge resources were spent for e.g. courseware development and not enough was left to improve the actual quality of learning [17]. Moreover, the interest received by user modelling aspect has not succeeded to address the variety and richness of the educational environment, even in the terms of user individual characteristics. These concerns have been ignored for quite long, hoping that new technology will somehow resolve the lack of real progress. The experience has proved so far that these issues cannot be wished away as they determine the type and scope of e-learning systems that are likely to succeed.

3.4 Pedagogical Framework

From a pedagogical point of view, an easy to use learning and teaching system is not necessarily also "properly designed". Namely, every e-learning experience begins with the pedagogy that drives it and continues with setting learning goals, designing learning activities which require the content suitable to meet the learning goals, *cf.* [15]. Although a consolidated pedagogical framework for e-learning does not exist thus far, it is argued that syllabus, course structure and design, navigation, assessment, collaboration and the like should be considered while defining framework components. E-learning system design based on selected pedagogical model would enable teachers to make use of the learning resources in a form which is appropriate to the learning goals and the particular learning style of the learner.

UNITE's pedagogical framework [32], covering theories and practices for enhancement of the learning experience with the use of ICT, is a good example of e-learning pedagogical basis design. This five-component pedagogical framework, through the combination of its constituents, produces the set of parameters that drive

and guide the creation of learning scenarios. UNITE's pedagogical framework components are:

- *pedagogical framework context*; defines areas that influence the framework itself and forms the basis for further development of UNITE's theories,
- *pedagogical approaches*; promote particularly principles of constructivist theory, along with blended, collaborative and active learning,
- *assessment techniques*; define and support diverse types of assessments as self-assessment, peer-assessment and so on,
- *teacher training*; supports teachers' work and endorse them during content production as well as delivery strategies decision and finally
- *current pedagogical practices* implemented in national curricula and national specifics.

Definition of the pedagogical model to be used as well as setting up pedagogical goals represents an integral part of the e-learning design process and should be detailed out well before technology choices and initial user interface design.

3.5 Guideline Framework

To facilitate design of usable interfaces and accessible systems in general, suitable guidelines and principles are essential. *Usability* should not only be evaluated, but "designed" from the start. Not only "the look" *i.e.* the visual design of the application, should be steered with proper guidelines, but also "the feel" *i.e.* the design of navigational and functional elements. Consequently, resulting e-learning application would be usable in view of a number of attributes: navigational fidelity, structure of hypermedia, system applicability, error handling, ease of use, consistency and standards, as well as performance evaluation, *cf.* [18].

Another system design issue that should be guided is *accessibility*. While some authors consider accessibility as one of usability's components, e.g. [34], others point out potential conflicts in including accessibility guidelines alongside usability ones, e.g. [22]. Most of the existing efforts for supporting the design, preparation and deployment of accessible e-learning propose guidelines that primarily address technical accessibility issues [6]. Although W3C has laid the foundation necessary for anyone to access and use web content [33], simple implementation of W3C guidelines does not suffice as they are not focusing on the needs of the users.

There should be strict guidelines for both usability and accessibility not only at the general level, but on more subsequent ones. Because any adaptation of the system once it is in place is time and resource consuming, adherence to proposed principles is advised. However, although applying good design guidelines by themselves is a good start, there is no substitute for e-learning system assessment.

4 Conclusion and Future Research Directions

In order to design e-learning systems with user interfaces which will not impede the effective learning and teaching, current HCI research is starting to address a learning perspective as well. Several recent studies to e-learning design are presented in the paper, offering a brief overview of diverse approaches to the identification of main

issues related to the inclusive design in the context of e-learning. These findings, along with our research and experience, provided a constructive input for the proposition of the approach which addresses main issues concerning design for all in e-learning environment. Such design, able to address the needs of all potential users, despite usability and learning perspective requires additional considerations, hence comprising: (i) learner-centred design paradigm, (ii) context of use approach, (iii) individualized approach, (iv) pedagogical framework and (v) guideline framework. Consequently, the main issues regarding universal design related to e-learning systems range from *learner-centred design* (thus taking into consideration *context of use* and *individual user characteristics*) to learning perspective or the *pedagogy* as a driving force of an e-learning experience as well as *guidelines* for good system design (related to usability and accessibility aspects). We propose addressing the distinctive needs of users as learners to augment the user-centred design framework. Special attention should be awarded to the context of e-learning usage defined as "a description of the actual conditions under which the e-learning system will be used in learning and teaching process". No less is to consider learners' individual characteristics as it is to define the pedagogical model and determine the pedagogical goals. Both learners' individual characteristics and the pedagogical model should be elaborated before making any technology choices or initial interface design. In order "to design for usability and accessibility", a set of strict guidelines on a general level and on a more subsequent one as well, has to be followed.

So far, our research has been focused on e-learning systems which mimic the capabilities of human teachers in the process of learning and teaching. The systems have been repeatedly evaluated concerning their efficiency in the learning process and usability of the interface design. Our future research will address issues concerning learning perspective, individual user characteristics, context of use and accessibility as well. The work related to the proposed approach is still in the early phase – a refinement from methodological point of view is needed along with the examples of its appliance. In order to build up our five-aspect approach, future research directions will embrace:

- specification of the final list of criteria per each aspect (with acceptable/recommended values and/or "best-practice" examples); the focus will be more on those aspects not yet addressed,
- development of the appropriate methodology; so far the combination of behaviour and opinion founded measures (linked scenario-based user testing with a kind of "less formal" heuristic evaluation, i.e. guideline evaluation) was applied, and
- an empirical study to test the efficiency of proposed approach will be fundamental.

A coherent universal framework, a support for the design of e-learning systems able to address the needs of learners while performing learning activities in their learning environment, will be provided.

Acknowledgments. This work has been carried out within project 177-0361994-1998 *Usability and Adaptivity of Interfaces for Intelligent Authoring Shells* funded by the Ministry of Science and Technology of the Republic of Croatia. Definition of Pedagogical framework presented in the paper has been carried out within the project 026964 *UNITE: Unified eLearning environment for the school*, partially supported by the European Community under the Information Society Technologies (IST) priority of the 6th Framework Programme for R&D.

References

1. American Society of Training and Development ASTD (2001) [Available On-line] http://www.astd.org/astd
2. Ardito, C., Costabile, M.F, De Marsico, M., Lanzilotti, R., Levialdi, S., Roselli, T., Rossano, V.: An approach to usability evaluation of e-learning applications. Universal Access in Information Society 4, 20–283 (2006)
3. Beaudouin-Lafon, M.: Designing Interaction, not Interfaces. In: Proceedings of the working conference on Advanced visual interfaces, AVI 2004, Gallipoli, Italy, pp. 15–22. ACM Press, New York (2004)
4. Benyon, D., Crerar, A., Wilkinson, S.: Individual Differences and Inclusive Design. In: Stephanidis, C. (ed.) User Interfaces for All: Concepts, Methods, and Tools, pp. 21–47. LEA, Mahwah, New Jersey (2001)
5. Costabile, M., Marisco, M., Lanzilotti, R., Plantamura, V., Roselli, T.: On the usability evaluation of E-learning applications. In: Proceedings of the 38th Annual Hawaii International Conference on System Sciences (HICSS'05) – Track 1 – vol. 01, 6.2 (2005)
6. De Marsico, M., Kimani, S., Mirabella, V., Norman, K.N., Catarci, T.: A proposal toward the development of accessible e-learning content by human involvement. Universal Access in Information Society 5, 150–169 (2006)
7. Dringus, L.: An iterative usability evaluation procedure for interactive online courses. Journal of Interactive Instruction Development 7, 10–14 (1995)
8. Egan, D.: Individual Differences in Human-Computer Interaction. In: Helander, M. (ed.) Handbook of Human-Computer Interaction, pp. 543–568. Elsevier Science B.V. Publishers, North-Holland (1988)
9. Good, J., Robertson, J.: CARSS: A Framework for Learner-Centred Design with Children. International Journal of Artificial Intelligence in Education 16(4), 381–413 (2006)
10. Granić, A.: Human-centred Design in Intelligent Tutoring: a Key Role of Usability Evaluation. In: Proceedings of the 3rd Cambridge Workshop on Universal Access and Assistive Technology, CWUATT 06, Fitzwilliam College, pp. 121–129. University of Cambridge, Cambridge, United Kingdom (2006)
11. Healey, D.: Theory and research: Autonomy in language learning. In: Egbert, J., Hanson-Smith, E. (eds.) CALL environments: Research, practice, and critical issues. Alexandria, VA: Teachers of English to Speakers of Other Languages, Inc. pp. 391–402 (1999)
12. Hsi, S., Soloway, E.: Learner-centred design: Addressing, finally, the unique needs of learners. In: Computer Human Interaction '98, CHI '98, Los Angeles, USA, pp. 211–212. ACM Press, New York (1998)
13. ISO Ergonomic requirements for office work with visual display terminals (VDTs) – Part 11: Guidance on usability (ISO 9241-11:1998(E)). Geneva, Switzerland
14. Keates, S., Clarkson, J.: Countering Design Exclusion: An Introduction to Inclusive Design. Springer, London (2003)
15. Kelly, B., Phipps, L., Swift, E.: Developing A Holistic Approach for E-Learning Accessibility. Canadian Journal of Learning and Technology, vol. 30(3) (2005)
16. Kinshuk, Patel, A, Russell, D.: Intelligent and adaptive systems. In: Collis, B., Adelsberger, H., Pawlowski, J. (eds.) Handbook on Information Technologies for Education and Training, pp. 79–92. Springer, Heidelberg (2001)
17. Nielsen, J.: Jacob Nielsen on e-learning. Elearningpost. (January 16th, 2001) [Available On-line] http://www.elearningpost.com/features/archives/001015.asp
18. Nielsen, J.: Heuristic evaluation. In: Nielsen, J., Mack, R.L. (eds.) Usability Inspection Methods, pp. 25–62. John Wiley& Sons, Inc, New York (1994)

19. Norman, D., Draper, S.W. (eds.): User centred system design. Erlbaum, Hillsdale NJ (1986)
20. Notess, M.: Usability, user experience and learner experience. eLearn Magazine. ACM Press, New York, NY, USA (2001) [Available On-line] http://www.elearnmag.org
21. Parlangeli, O., Marchigiani, E., Bagnara, S.: Multimedia systems in distance education: effects of usability on learning. Interacting with Computers 12, 37–49 (1999)
22. Phipps, L., Kelly, B.: Holistic approaches to e-learning accessibility. Association for Learning Technology 14(1), 69–78 (2006)
23. Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S., Carey, T.: Human-Computer Interaction. Addison-Wesley, Wokingham, England (1994)
24. Schneiderman, B., Plaisant, C.: Designing the User Interface. Strategies for Effective Human-Computer Interaction, 4th edn. Addison-Wesley, Reading, MA (2005)
25. Sloan, D., Kelly, B., Heath, A., Petrie, H., Fraser, H., Phipps, L.: Contextual Web Accessibility - Maximizing the Benefit of Accessibility Guidelines. In: Conference Proceedings of WWW, Edinburgh, Scotland May 2006. Special Interest Tracks, Posters and Workshops (2006) [Available On-line] http://www.ukoln.ac.uk/web-focus/papers/w4a-2006/
26. Soloway, E., Jackson, S., Klein, J., Quintana, C., Reed, J., Spitulnik, J., Stratford, S., Studer, S., Jul, S., Eng, J., Scala, N.: Learning Theory in Practice: Case Studies of Learner-Centred Design. In: Proceedings of Conference on Human Factors in Computing Systems (CHI96), Vancouver, British Columbia, pp. 189–196. ACM Press, New York (1996)
27. Soloway, E., Guzdial, M., Hay, K.E.: Learner-centred design: the challenge for HCI in the 21st Century. Interactions 1, 36–48 (1994)
28. Squires, D., Preece, J.: Predicting quality in educational software: Evaluating for learning, usability and the synergy between them. Interacting with Computers 11, 467–483 (1999)
29. Squires, D., Preece, J.: Usability and Learning: Evaluating the Potential of Educational Software. Computers Education 27(1), 15–22 (1996)
30. Stephanidis, C.: User Interfaces for all: New Perspectives into Human-Computer Interaction. In: Stephanidis, C. (ed.) User Interfaces for All: Concepts, Methods, and Tools, pp. 3–21. LEA, Mahwah, New Jersey (2001)
31. Stephanidis, C., Salvendy, G., Akoumianakis, D., Bevan, N., Brewer, J., Emiliani, P.L., Galetsas, A., Haataja, S., Iakovidis, I., Jacko, J., Jenkins, P., Karshmer, A., Korn, P., Marcus, A., Murphy, H., Stary, C., Vanderheiden, G., Weber, G., Ziegler, J.: Towards an Information Society for All: An International R&D Agenda. International Journal of Human-Computer Interaction 10(2), 107–134 (1998)
32. UNITE FP6 IST project. Contract No. 026964. Website http://www.unite-ist.org
33. Web Content Accessibility Guidelines 2.0. W3C Working Draft (April 27, 2006) [Available On-line] http://www.w3.org/TR/WCAG20/
34. Zaharias, P.: E-learning design quality: A holistic conceptual framework. In: Howard, C., Boettcher, J., Justice, L., Schenk, K., Rogers, P.L., Berg, G.A. (eds.) Encyclopedia of Distance Learning, vol. II, Idea Publishing, New York (2005)
35. Zaharias, P., Vassilopoulou, K., Poulymenakou, A.: Designing on-line learning courses: Implications for usability. Scientific Journal on Applied Information Technology vol. 1(1) (2002)

Designing Intelligent Interfaces for e-Learning Systems: The Role of User Individual Characteristics

Andrina Granić and Jelena Nakić

University of Split, Faculty of Natural Sciences, Mathematics and Education
Nikole Tesle 12, 21000 Split, Croatia
{andrina.granic,jelena.nakic}@pmfst.hr

Abstract. In order to advance personal learning experience it is crucial to overcome the one-size-fits-all approach in user interface design and increase the e-learning systems intelligent behavior. Recent research has confirmed that user individual characteristics must be taken into account to accomplish that goal. This paper identifies user features relevant for system's adaptation in general. Aiming to investigate affect of those features on users' learning outcomes in e-learning environment in particular, an empirical study along with obtained results is reported as well.

Keywords: intelligent user interfaces, adaptive user interface, e-learning, individual differences.

1 Introduction

"We have spent the last 50 years learning how to create computers and how to program computers. The next frontier is in actually making computers serve and adapt to human needs rather than forcing humans to adapt." claimed Dan Olsen in the already remote ACM UIST 97 panel on the future of human-computer interaction (HCI) [13, p. 118]. In order to achieve a better fit to this agenda, recent HCI research with its user-centered design [18] and learner-centered design [24] approaches puts the user, the individual at the centre of all developments, stressing the importance to design technologies for human needs. Consequently, the role of an intuitive user interface and a flexible interaction suited to different needs becomes even more important for the users' success, as users with a wide variety of background, skills, interests, expertise and goals are using computers for quite diverse purposes. In such context, the role of intelligent user interfaces (IUI) is unquestionable. IUIs should facilitate a more natural interaction between users and computers, not attempting to imitate human-human communication, but instead aiding the human-computer interaction process [12]. User modeling, as one of the employed techniques in IUI area, is used to denote a model of the user that the system maintains and adapts its behavior to. As the range and complexity of interactive system increases, understanding how a system can dynamically capture relevant user needs and features, and subsequently adapt its interaction, has become vital for designing effective interfaces.

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 627–636, 2007.
© Springer-Verlag Berlin Heidelberg 2007

Focusing on e-learning, an instructional content or learning experience delivered or enabled by electronic technology [19], despite so much publicity and activity, the progress in the field has been unexpectedly slow. In order to improve the learning experience and increase e-learning system's intelligent behavior, respective interaction mechanisms merit additional consideration. Apparently, there should be a synergy between the learning process and a user's/learner's interaction with the e-learning application [25], additionally taking into account the different ways user's learn and ensuring their natural and flexible interactions as well. Most of the current e-learning applications are static and inflexible, designed without considering users' preferences and abilities. It is vital to overcome the one-size-fits-all approach and provide users with individual learning experiences through e-learning systems with intelligent user interface.

Our research has been focused on the employment of intelligence in e-learning system's interface with the intention of adjusting the system to individual users. In order to design adaptive interaction in e-learning applications, our recent work includes an application of adaptation techniques, along with methodology of user modeling, from the field of adaptive hypermedia (AH) [9]. To enhance and to broaden this AH based approach, we have analyzed user features revealing personal differences relevant for system intelligent behavior in the HCI field in general, as well as identified and quantified those we find significant for e-learning system adaptation in particular. In this paper we concisely inform about those features in both domains and then report the role of individual user characteristics in design of intelligent user interfaces for e-learning systems. The main objective of the presented research study was to examine the affect of users' individual differences on their learning outcomes achieved in e-learning environment. Obtained results will enable a more precise design of adaptation technology in e-learning systems.

The rest of the paper is structured as follows. Section 2 brigs taxonomy of user features revealing individual differences and stresses the importance of intelligent systems to adapt individual users, especially in educational environment. The empirical study presented and discussed in Section 3 aimed to investigate the role of users' personal characteristics in their knowledge acquisition process while using e-leaning application. Section 4 summarizes achieved results and concludes the paper.

2 Background

IUIs have been recommended as a means for making systems individualized or personalized, thus enhancing the systems flexibility and attractiveness [12]. The intelligence in interface can for example make the system adapt to the needs of different users, take initiative and make suggestions to the user, learn new concepts and techniques or provide explanation of its actions *cf.* [12], [15].

The research in HCI field has already confirmed and empirically proved that system intelligent behavior strongly relies on individual differences [6], [7], [8], [14], [16], and on the other hand has implications on the degree of success or failure experienced by users. When considering adaptation of system to individual use, user personality and cognitive factors have to be taken into account because of their higher resistance to change. Moreover, it is useful to exploit a certain amount of "stable"

knowledge about the user conveyed through long-term characteristics, containing information about user's level of expertise with computers in general, her/his expertise with the system in particular, as well as familiarity with the system's underlying task domain. Certain information related to user's preferences or current goals conveyed through short-term user characteristics should also be taken into account. Table 1 provides taxonomy of key user characteristics for system adjustment and brings in several authors who consider them important for user description as well. Those features are generally categorized as:

- *personal user characteristics*, quite stable over time and independent from the system, where (i) general personal characteristics (including characteristics that reflect internal psychological state) and (ii) previously acquired knowledge and abilities can be differentiated, as well as
- *system-dependent user characteristics*; the most changeable category of characteristics as related to particular system.

Table 1. User characteristics revealing individual differences; A-Benyon and Murray [2], B-Egan [7], C-Browne *et al.* [3], D-Norico and Stanley [17], E-Dillon and Watson [6], F-Rothrock *et al.* [23]

		A	*B*	*C*	*D*	*E*	*F*
personal characteristics	gender	●	●			●	●
	age		●			●	
	personality & affections	●	●	●		●	●
previously acquired knowledge and abilities	experience	●	●	●	●	●	●
	cognitive abilities	●			●	●	●
	psycho-motor skills			●		●	●
	technical aptitudes		●		●		
	domain knowledge	●	●	●	●	●	●
system dependent characteristics	goals & requirements			●		●	●
	motivation			●		●	●
	expectations			●			●

Yet, when considering e-learning systems, it has been claimed that "… although technology is often touted as the great salvation of education – an easy way to customize learning to individual needs – it rarely lives up to this broad expectation" [11, p. 398]. Interest received by user modeling aspect has not succeeded to address the variety and richness of the educational environment, even in the terms of user individual characteristics. Learners have different requirements like their individual learning style, their actual learning in the learning process and their individual background knowledge. These issues have been ignored for quite a long, hoping that new technology will somehow resolve the lack of real progress. The experience has proved so far that these issues cannot be wished away as they determine the type and scope of e-learning systems that are likely to succeed [20].

3 Research Method

The experimental study reported in the following aimed to question existence and level of interaction between users' individual differences and learning outcomes achieved while using an e-learning application. Personal user features assumed to affect learning process were clearly identified and the methods how to measure them were determined. We have classified characteristics intended to be measured in two categories: (i) intelligence and personality factors, and (ii) experience, motivation, expectations and background knowledge.

3.1 Subjects

Undergraduate students of the fourth semester from two different university programs (mathematics–computer science and physics–computer science) were recruited to the briefing where they have been introduced with nature and purpose of the experiment. Since we intended to use an e-learning application on programming domain in this study, we have consequently selected twenty-four students (6 males and 18 females) among volunteers who hadn't taken an exam on Programming I before. The participants of the study have been told that their achievement on test would have only experimental use and would not affect their future exam grades.

3.2 Research Instruments

Standard psychological tests were used to measure subjects' personal characteristics. Intelligence test (D-48) measured general mental abilities, while personality test (EPQ) measured dimension of emotional stability/instability, extraversion/introversion, mental stability/psychoticism and honesty/dissimulation level [21]. Those tests are protected by national laws and regulations so cannot be presented here.

A questionnaire was developed in order to obtain students' gender, experience in using computers and Internet, motivation and expectations from e-learning. Students' grades from Introduction to Computer Science exam were taken as a measure of their background knowledge to material supposed to be learned during the experimental session.

E-learning application used to test students' knowledge is an intelligent tutoring system developed and evaluated on our faculty [26]. Intelligent tutoring systems (ITSs) are computer-based educational systems which emulate human teacher in order to support the process of learning and teaching in arbitrary domain knowledge [4]. We consider this Web-based application as well-accepted instrument for this research since its effectiveness has been evaluated in several case studies, e.g. [10], and it has been shown that system can support at least 20 users at a time.

Participants of the experiment were already familiar with functionality of the ITS since similar systems (generated from the same authoring shell [26]) have been used in university courses in the first and second semester. However, the students have never had access to learning modules or quiz related to course Programming I, which were selected to facilitate in this study.

3.3 Procedure

Experiment was conducted through five steps as shown in Fig. 1. First, a physiologist and a HCI professor interviewed the experimental group of students in order to introspect some general characteristics of the group so they could design a questionnaire related to their experience in using computers and Internet, motivation and expectations from e-learning and concrete system. Few days after the introductory interview the participants were invited to take intelligence test and personality test.

Two experimental sessions in an on-line classroom were conducted for groups of 12 students at a time. Students were not allowed to take notes or use any external learning material, paper or on-line, besides the system lessons on selected ITS. They were free to learn for 30 minutes, and then began to test acquired knowledge on a quiz belonging to particular ITS simultaneously. Time for testing was limited to 15 minutes and all participants completed the quiz at given time.

After completing the quiz, students were asked to fill in the multiple choice questionnaire prepared to gain data about their gender, prior experience in using computer and Internet, motivation to learn programming, expectations from e-learning systems in general and also expectations and satisfaction with an ITS they have just exploited.

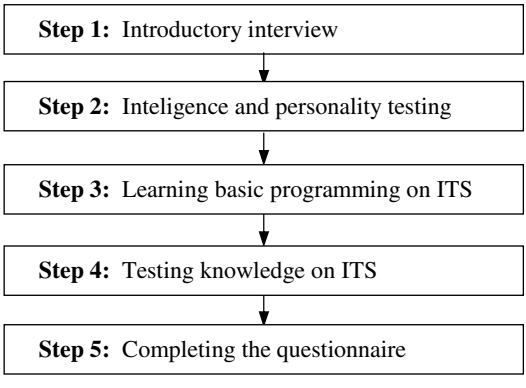


Fig. 1. Five-steps procedure of the experiment

3.4 Results

Data analysis of 24 complete datasets was performed using Statistica 7 software package. All variables (except gender) were measured on interval scales and relations between variables are expressed by Pearson correlations [22]. Significance level of $p < 0.05$ was considered acceptable for this research. We ensured that obtained values of all relevant variables were normally distributed (e.g. quiz results have Chi-Square test = 1.16499, df = 1 (adjusted), p = 0.28043). Since 24 participants constitute a small sample, the sample has been analyzed as a whole, without splitting at the mean by any variable.

Although the main issue of the research was to investigate the influence of user individual differences on their learning outcomes accomplished on e-learning system, it was inspiring for us to see if there were any connections among individual differences of the users themselves. Some interesting correlations were found between intelligence and personality factors obtained by tests with other user characteristics obtained by questionnaire. Those results are given in Table 2. Significant correlations were found between mental stability and motivation ($r = -0.50$, $p < 0.05$) and also between emotional stability and expectations from using the system ($r = -0.45$, $p < 0.05$). This means that mentally stable students are more motivated to learn programming then mentally unstable or "more neurotic" students. Similar to that, emotionally stable students have greater expectations from e-learning than emotionally unstable or "more psychotic" students. Highly significant correlation was found between students' prior experience in using computers and Internet and their background knowledge required to learn programming ($r = 0.62$, $p < 0.01$) as expected. It seems that students' intelligence and dimension of extraversion/introversion are not associated with any other individual characteristics.

Table 2. Pearson correlations between user individual characteristics

	Intelligence	Emotional stability	Extraversion	Mental stability	Experience
Experience	0.22	-0.28	0.19	-0.18	
Motivation	-0.17	-0.08	0.07	-0.50*	0.16
Expectations	-0.11	-0.45*	0.36	-0.26	0.01
Background	0.39	-0.28	0.06	-0.11	0.62**

*Significant correlations at level of $p < 0.05$.
** Significant correlations at level of $p < 0.01$.

Table 3 shows correlations between students' individual differences and their learning outcomes accomplished while using e-learning application. Apparently there are no associations between intelligence and personality factors with learning outcomes. Considering characteristics obtained by questionnaire, it seems that only motivation to learn programming in addition to expectations of e-learning has

Table 3. Pearson correlations between user individual characteristics and knowledge acquired on e-learning system

	Intelligence	Emotional stability	Extraversion	Mental stability	Experience
Knowledge	0.05	-0.29	-0.00	-0.15	0.29

	Motivation and expectations	Background
Knowledge	0.42*	0.26

* Significant correlations at level of $p < 0.05$.

statistically significant impact on knowledge acquired through interaction with the system ($r = 0.42$, $p < 0.05$).

Analysis by age and by prior experience in using concrete system was not conducted because individual differences among participants were minor in those variables. Also, analysis by gender would be ineffectual because the subjects group consisted of 6 males and 18 females, which are small samples.

3.5 Discussion

There are numerous studies reporting minor influence of personality factors on predictions of user performance (reviewed in [6]) or no influence at all [7], so the perceived lack of associations between intelligence and personality with learning outcomes in our analysis was not quite unexpected.

However, careful interpretation and observation of the obtained results revealed some shortcomings of applied methodology. First of all, the sample we analyzed was too small and too homogenous to give us strong grounds for generalization of the results. All participants of the experiment were students of the same age, with comparable background knowledge and experiences, intellectual capabilities and motivation for graduating as well. Similar experiment with larger sample of more diverse users would certainly provide more reliable results. Thus we consider this study as a pilot experiment that gave us important directions to establish an enhanced methodology for future research.

When we observe the connection between subjects' experience in using computers and Internet with their background knowledge related to programming (see Table 2), the lack of similar connection with their quiz scores ($r= 0.26$ as shown in Table 3) was quite unexpected. This result is more surprising in comparison to related work firmly confirming the major affect of experience on task outcomes e.g. [5], [7], [8], [17]. Part of the explanation of that result could be found in insufficient diversity of students' experience. Besides that, we cannot ignore a procedural issue that possibly caused more homogenous scale of quiz results then expected. The ITS on programming domain was selected for the research because none of the subjects hadn't taken an exam on Programming I before participating in experiment and knowledge tested on the ITS was accepted as reliable variable for the experiment [10]. Still, many of the students participated in experiment had attended classes related to programming which could have increased their background knowledge on particular domain. One possible solution of that situation is to select an ITS on some other domain knowledge, completely unknown to the subjects. Even better solution is to give them a pre-test and post-test (best the same as pre-test) from the same domain knowledge we intended to measure. This could be paper-and-pen test or simply the quiz on particular ITS, given before and after the learning session. The gain between those two tests scores would be considered as learning outcome achieved on e-learning system only through that particular session. In addition, the time required to complete the test could be used as another aspect of learning outcome for each participant (as suggested e.g. in [1], [7]).

Since a number of participants achieved relatively high dissimulation level on personality test it is possible that they have unrealistically high opinion on their own knowledge. According to that assumption we computed correlation between students'

answers on two questions measuring their expectations of e-learning (see Table 4) but found no actual connection between those answers. This result is in contrary to empiric hypothesis that students who consider e-learning effective would be satisfied with their learning outcomes. Possible explanation is that a question positioned between those two and regarding their grades achieved on quiz, affected their answers on following question. Although we cannot affect the subjects' honesty/dissimulation level, more careful composing and reliability analysis of requisite questionnaire, as well as increasing number of questions regarding particular variable would obviously lead to more accurate measurements of user characteristics.

Table 4. Correlation between two questions regarding users' expectations of e-learning

Question no. 7.	You consider e-learning generally a) very efficient and necessary nowadays b) efficient but not necessary c) medially efficient d) inefficient e) inefficient and distressing
Question no. 9.	You think your grade on the quiz is a) completely accurate measure of knowledge b) very accurate measure of knowledge c) half-accurate measure of knowledge d) barely measure of knowledge e) no measure of knowledge at all
r = 0.39	

In this study we did not investigate the affect of experience in using concrete e-learning application on learning outcomes on the same application. This could be done by comparing the learning outcomes of two independent groups of subjects, one experienced in using similar ITSs (generated from the same authoring shell but for different domain) and the other without such experience.

4 Conclusion

Intelligent user interfaces should facilitate a more natural communication between users and computers, aiding the human-computer interaction process. Our research has been focused on the employment of intelligence in e-learning system's interface in order to adjust the system to individual users. In account to that we have analyzed and reported user features revealing personal differences relevant for system intelligent behavior in the HCI field in general, and in educational domain in particular.

This paper also presents the empirical study we have conducted in order to examine the influence of users' individual differences on their knowledge acquisition process in e-leaning environment. We have analyzed interrelations among quantified personal characteristics and found positive associations of mental stability with motivation and emotional stability with expectations from e-learning. Highly

significant correlation was discovered between students' prior experience in using computers and Internet with their background knowledge, but similar connection of experience and learning outcomes achieved on the e-learning application was not found. This experiment indicated that motivation to learn programming in addition to expectations of e-learning significantly affects on users' learning achievement.

The observed sample had certain limitations, especially considering the number and diversity of subjects who were all students of the same age. Aware of the great sensitivity of results to the sample, instead of generalization of presented results we have used them to determine the guidelines for developing further research design. Such research is clearly needed to be conducted in order to provide us stronger foundations for designing adaptation mechanism of e-learning systems.

Acknowledgments. This work has been carried out within project 177-0361994-1998 *Usability and Adaptivity of Interfaces for Intelligent Authoring Shells* funded by the Ministry of Science and Technology of the Republic of Croatia. The authors would like to thank psychologist assoc. prof. M. Nazor and graduate student I. Đolonga for their valuable contribution in conducting experimental part of the study.

References

1. Benyon, D., Höök, K.: Navigation in Information Spaces: Supporting the Individual. INTERACT 1997 pp. 39–46 (1997)
2. Benyon, D., Murray, D.: Developing Adaptive Systems to Fit Individual Aptitudes. In: Gray, W.D., Hefley, W., Murray, D. (eds.) International Workshop on Intelligent User Interfaces 1993 (January 4-7, 1993) pp. 115–12 Orlando, Florida, USA (1993)
3. Browne, D., Norman, M., Rithes, D.: Why Build Adaptive Systems? In: Browne, D., Totterdell, P., Norman, M. (eds.) Adaptive User Interfaces, pp. 15–59. Academic Press, London (1990)
4. Burns, H., Capps, C.: Foundations of Intelligent Tutoring Systems: An Introduction. In: Polson, M., Richardson, J. (eds.) Foundations of Intelligent Tutoring Systems, pp. 1–18. Lawrence Erlbaum Associates Publishers, Hillsdale, NJ (1998)
5. Chen, C.: Individual Differences in a Spatial-Semantic Virtual Enviroment. Journal of the American Society for Information Science 51(6), 529–542 (2001)
6. Dillon, A., Watson, C.: User Analysis in HCI – The Historical Lessons From Individual Differences Research. International Journal on Human-Computer Studies 45, 619–637 (1996)
7. Egan, D.: Individual Differences in Human-Computer Interaction. In: Helander, M. (ed.) Handbook of Human-Computer Interaction, pp. 543–568. Elsevier Science B.V. Publishers, North-Holland (1988)
8. Ford, N., Chen, S.Y.: Individual Differences, Hypermedia Navigation and Learning: An Empirical Study. Journal of Educational Multimedia and Hypermedia 9(4), 281–311 (2000)
9. Granić, A., Stankov, S., Nakić, J.: Designing Intelligent Tutors to Adapt Individual Interaction. In: 9th ERCIM Workshop User Interfaces for All (UI4All) – Workshop UI4All 2006 Proceedings (September 27-28, 2006) Konigswinter (Bonn), Germany (in press)

10. Grubišić, A., Stankov, S., Žitko, B.: An Approach to Automatic Evaluation of Educational Influence. In: Proceedings of the 6th WSEAS International Conference on Distance Learning and Web Engineering (DIWEB '06), Lisbon, Portugal (September 22-24, 2006) pp. 20–25 (2006)
11. Healey, D.: Theory and Research: Autonomy in Language Learning. In: Egbert, J., Hanson-Smith, E. (eds.) CALL Environments: Research, Practice, and Critical Issues. Alexandria, VA: Teachers of English to Speakers of Other Languages, Inc. pp. 391–402 (1999)
12. Hook, K.: Steps to Take Before Intelligent User Interfaces Become Real. Journal of Interaction with Computers 12(4), 409–426 (2000)
13. Jacob, R., Feiner, S., Foley, L., Mackinalay, J., Olsen, D.: UIST'007: Where Will We Be Ten Years From Now? In: Proceedings of ACM UIST'97 Symposium on User Interface Software and Technology, pp. 115–118. Addison-Wesley/ACM Press, Banff, Canada, Panel statement (1997)
14. Jennings, F., Benyon, D., Murray, D.: Adapting systems to differences between individuals. Acta. Psychologica 78, 243–256 (1991)
15. Lieberman, H.: Introduction to intelligent interfaces. (1997) [Available On-line] http://web.media.mit.edu/˜lieber/Teaching/Int-Int/Int-Int-Intro.html
16. Magoulas, G.D., Chen, S.Y., Papanikolaou, K.A.: Integrating Layered and Heuristic Evaluation for Adaptive Learning Environments. In: Weibelzahl, S., Paramythis, A. (eds.) Proceedings of the Second Workshop on Empirical Evaluation of Adaptive Systems, held at the 9th International Conference on User Modeling UM2003, Pittsburgh pp. 5–14 (2003)
17. Norcio, A., Stanley, J.: Adaptive Human-Computer Interfaces: A Literature Survey and Perspective. IEEE Transactions on System, Man and Cybernetics 19(2), 399–408 (1989)
18. Norman, D., Draper, S.W. (eds.): User Centred System Design. Erlbaum, Hillsdale NJ (1986)
19. Pantazis, C.: Executive summary: A vision of E-learning for America's workforce. Report of the Commission on Technology and Adult Learning, ASTD (2001) [Available On-line] http://www.learningcircuits.org/2001/aug2001/pantazis.html
20. Patel, A., Kinshuk: Knowledge Characteristics: Reconsidering the Design of Intelligent Tutoring Systems. In: Behrooz, A. (ed.) Knowledge Transfer. pAce, London, pp. 190–197 (1996)
21. Petz, B. et al.: Psychological Dictionary (Psihologijski rječnik). Naklada Slap, Jastrebarsko (2005)
22. Petz, B.: Basic Statistical Methods for Non-mathematicians (Osnovne statističke metode za nematematičare). Naklada Slap, Jastrebarsko (2004)
23. Rothrock, L., Koubek, R., Fuchsm, F., Haas, M., Salvendy, G.: Review and Reappraisal of Adaptive Interfaces: Toward Biologically Inspired Paradigms. Theoretical Issues in Ergonomics Science. pp. 47–84 Taylor and Francis Ltd, vol. 3(1) (2002)
24. Soloway, E., Guzdial, M., Hay, K.E.: Learner-Centred Design: The Challenge for HCI in the 21st Century. Interactions 1, 36–48 (1994)
25. Squires, D., Preece, J.: Usability and Learning: Evaluating the Potential of Educational Software. Computers Education 27(1), 15–22 (1996)
26. Stankov, S.: Principal Investigating Project TP-02/0177-01 Web-oriented Intelligent Hypermedial Authoring Shell. Ministry of Science and Technology of the Republic of Croatia (2003–2005)

Providing Universal Accessibility Using Connecting Ontologies: A Holistic Approach

Shuaib Karim[1,2], Khalid Latif[1,3], and A. Min Tjoa[1]

[1] Institute of Software Technology & Interactive Systems, TU Wien - Austria
[2] Department of Computer Science, Quaid-i-Azam University, Islamabad – Pakistan
[3] Department of Computer Science & I.T., Islamia University, Bahawalpur – Pakistan
{skarim,klatif,amin}@ifs.tuwien.ac.at

Abstract. Accessibility implementation efforts are generally focused towards some typical user impairments and a few interaction devices for a particular user scenario. Whereas there are numerous factors in the prevalent context which can leverage the overall accessibility to its full potential, only if those are orchestrated together. In absence of a common and unifying approach the industry has little choice in abstaining from legacy and is therefore not very successful in producing universally accessible software. Exploiting enabling technology of Semantic Web, we present an approach by developing Connecting Ontologies for these different factors. Exemplars are developed to show its correctness and practicability.

Keywords: Universal Accessibility, Diversity, Connecting Ontologies, Mapping Ontologies, Ontology Design Patterns.

1 Introduction

In absence of a generic framework the provision of universally accessible software is a precarious task for the producers. Attempts for independent resolution of accessibility and diversity issues, which indeed complement each other, have further elongated the goals for both and the vague boundary between the two still prevails. According to descriptions by W3C[1], ANSI[2] and ISO[3] accessibility is not only related with user impairments but a wider range of contributing factors which are part of the overall context. Different types of interaction devices, varying user needs & impairments, and the specifications of the task at hand are some significant components of the context. User interfaces and visualizations are the means to carry out the tasks in a particular context. Also, visualizations are not *a priori* suitable for all types of information entities which are a combination of data and the related semantics for their description. Consequently, there are many contributing components interacting with each other. **In order to improve universal accessibility the components' semantics must be exploited to produce an accessible presentation for the end user. In this paper special attention is laid on accessibility for users with special needs or impairments.**

[1] http://www.w3.org/TR/WAIWEBCONTENT/ (January 15, 2007)
[2] http://acm.org/sigchi/bulletin/1997.2/standards.html#HDR3 (January 15, 2007)
[3] Available from http://www.stcsig.org/sn/internetISO.shtml (January 15, 2007)

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 637–646, 2007.
© Springer-Verlag Berlin Heidelberg 2007

Automation of this process consists of two major steps; a) formal description of semantics for each component and b) formal description of semantics of consequences and effects of potentially interacting component on each other. Ontologies for the formal semantic description of user impairments, device profiles, tasks related to a particular domain, and their inter-connections are being developed using recently emerging semantic web technologies. Ontologies could be developed and arranged by following different approaches such as taxonomic and faceted. For this research we have restricted ourselves to a hierarchical approach which will lead to further approaches in future. Instead of developing the ontology from scratch we have adopted concepts mainly from existing specifications, domain thesauri and foundational ontologies. In this paper we propose a framework for providing accessibility at a generic level using the notion of the *Connecting Ontologies.*

This work is an extension of our previous paper [1] which describes the connection between user interface characteristics and the user's impairments data using OWL-DL. The paper is structured as follows:

Next section describes the related work. Section 3 describes the concept of *Connecting Ontology* and how it is different from other apparently similar concepts like ontology mapping and ontology integration. This is followed by our methodology of exploiting ontology design patterns in Section 4, leading to architecture of Accessibility Framework in Section 5. The description of the method adapted for connecting user's life time information items & events is presented in Section 6. Finally, conclusions and future directions are given.

2 Related Work

The usefulness of connected knowledge was highlighted in [2] by describing the initial results where medical domain ontology can be seamlessly scaled and integrated with ontology of another domain. Similar discussions lead to the foundation of *Semantic Web Health Care and Life Sciences Interest Group*[4]. However, to the best of our knowledge the concept of connecting ontologies is not yet fully exploited to connect heterogeneous domains, especially for improving accessibility.

In [3] the concept is used to connect external business processes and internal work flow processes using *LAV (Local-As-View)* data integration approach [4] to map the two models. *LAV* approach provides a uniform query interface using a global mediated schema to be defined independent of the data sources. Its relationship with the data sources is then made possible by connecting global schema with specialized views for each data source. Looking at the *Accessibility Framework* (see Fig. 4) there is a conceptual similarity with our approach where *Query Interface* is realized using a global schema and the *Connecting Ontologies* are acting as mediation interface between global schema and the data sources (ontologies and instances for context components). However their concern is to connect process work flow models pertaining to "similar domain", which is not the case in our work.

The recent work described in [5] is an example of how OWL [6] in combination with RDQL rules [7] is employed for connecting heterogeneous ontologies in an

[4] http://www.w3.org/2001/sw/hcls/ (January 15, 2007)

electronic medical record application. More specifically, the concepts in SNOMED[5] are linked with an ontology containing drug / medicine related concepts such as drug classes, their interactions, allergies and formularies. The purpose is to limit the generation of disputed medical insurance bills which are created due to inconsistencies in coding schemes used during diagnosis (such as ICD9CM[6]) and the corresponding list of approved medical procedures as permitted by insurance companies. The semantic annotations are applied in XML files which allow using Web technologies such as XSL and XPATH, and rules interpretation by RDQL. However, the domain is not modeled in OWL which confines the benefits because of inability to use convenient querying by SPARQL and sophisticated inference capabilities offered by ontology reasoners.

The work about Contextualizing Ontologies [8] shows the mappings of GALEN medical ontology with Tambis genetic ontology by aligning both with another medical ontology called UMLS. Bridging rules using C-OWL (Context-OWL), are defined for mapping the individual concepts or concept expressions belonging to ontologies in similar domain of discourse. In contrast, our approach provides interconnection between concepts in varying domains of discourse.

3 Connecting Ontology

An ontology formally describes the concepts in the domain of discourse [9] or more realistically speaking, helps to formally specify the concepts. A System, by definition an integrated whole, is essentially composed of heterogeneous components or domains which have to interact with each other to achieve a certain goal. Each of these has its domain ontology which is composed of concepts description about that domain. By *Connecting Ontology* we mean an ontology which links two heterogeneous ontologies or in other words, which describes the linking of heterogeneous entities (concepts, relations and properties) across two ontologies. This is analogous to the famous wine and food ontology described in [10].

To illustrate it further, let us assume that ontology O_1 has a concept c_1, and ontology O_2 has another concept c_2. The *Connecting Ontology Oc* will contain either the concept c_3 to link c_1 with c_2 or extend c_1 with new features which in turn links it with c_2. Once the ontologies O_1 and O_2 are populated with instances, their connections are dynamically generated by the reasoner based on their connections present in *Oc*. It is often the case that the two ontologies were developed based upon different formalisms ensuring decidability of axioms. Although components of both the ontologies may be decidable, but when combined together the decidability may not be guaranteed. As described in [11], E-Connection is such a method to link two ontologies in terms of Abstract Description Systems (ADS) [12].

The concept of *Connecting Ontology* is different from other apparently similar notions like ontology mapping, alignment, articulation, merging or integration. The term *ontology mapping* is described as "the task of relating the vocabularies of two ontologies that share the same domain of discourse in such a way that the mathematical structure of ontological signatures and their intended interpretations, as

[5] http://www.snomed.org/snomedct/index.html (January 15, 2007)
[6] http://icd9cm.chrisendres.com/ (January 15, 2007)

specified by the ontological axioms, are respected" [13]. *Ontology alignment* is concerned with the process where binary relations (also called atomic roles such as properties) between vocabularies (more specifically concepts or unary predicates within these vocabularies) of two ontologies belonging to the same domain of discourse are established. This leads to *articulation of two ontologies* when these binary relations are specified in terms of ontology in itself. Properly specified articulation ontology helps in *merging or fusion* of the concerned ontologies. *Ontology merging or fusion* is closer to our notion of *connecting ontologies.* However, contrary to fusion the two ontologies are not merged but connected using additional features and relations. The aspects of composition of ontologies to build a new unified ontology, extending the existing ontologies to build new ones, and incorporating ontologies into the applications are described under *ontology integration.* We envision the distinctive features of *Connecting Ontology* as follows:

- The ontologies to be connected are not related with the same domain of discourse. Though they may be part of the overall application.
- Their vocabularies are independent of each other. Even if there are apparent similarities, they are still assumed to be independent.
- Each of the ontologies to be connected is supposedly developed using its own design pattern(s) - explained later in Section 4.
- There might be similarity between the design patterns of two ontologies which may prove helpful for connecting them together.
- Connecting ontologies from heterogeneous domains are in fact creating new knowledge in the light of user's experience, whereas other notions of ontology mapping help to reorganize the existing knowledge.

Following are some of the benefits of *Connecting Ontologies*:

- Users think in an application or need oriented way, find available artifacts & then try connecting them together to fulfill their needs. *Connecting Ontologies* would facilitate this top-down approach.
- The incompatibilities between two ontologies are solved at the ontological level without delving into the application code.
- By elevating one ontology with rich design patterns based upon new user needs, it is possible to trace the corresponding effects on the other ontology due to already established *cause-effect* relationship between them.
- In other words, connecting ontologies are helping to automate the coding process.
- The approach will be beneficial for many domains including future interfaces for e-Learning. Travel on the interface and way finding consisting of tasks such as exploration, search and maneuvering can be greatly improved by interconnecting ontologies of user's tasks, impairments, interaction devices & visualizations in use.

4 Exploiting Ontology Design Patterns

Recent research in *Ontology Design Patterns* [14] is the stimulating factor for improvising some useful patterns for our system. *ODPs* help to identify the objectives and conceptualize scope and components of the application and the related ontologies.

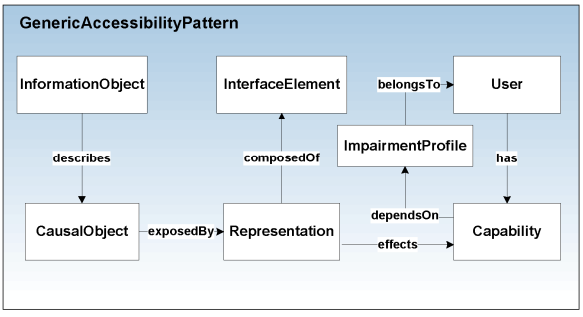


Fig. 1. A Generic Accessibility Pattern

4.1 Generic Accessibility Pattern

A generic accessibility pattern for connecting user profile and visualization ontologies is shown in Fig. 1. This pattern is further extended to implement various scenarios in *SemanticLIFE* system [15]. Its different components are described below. A few specializations are also shown which are explained later in this section.

- *InformationObject*: Information items in our system with associated semantics.
- *CausalObject*: Derived from information objects. For example, life events may be derived from the information items.
- *Representation*: The interface or visualization which is used for the visualization of the causal objects. There can be different visualizations for diverse users (specifically user's impairments profile contain impairments related data for people with special needs) in different contexts. The representation semantics are described in terms of their composition, intended users, data to be shown, & tasks for which it is designed. For example, tabular view is better for numerical data, whereas geographical map is more suitable for spatially significant information.
- *InterfaceElements*: Appropriate interface elements are selected for composing a representation based upon information objects and user characteristics. The semantics of interface elements are related with their usability measures and composition with other interface elements to form composite interface elements.
- *Capability*: It is the ability of the user to carry out a certain task. It depends upon the user's impairments data and the domain oriented task ontology. Based upon the user's task completion statistics the system can be fine-tuned by configuring the different components in the specified pattern.
- *ImpairmentsProfile*: This is the ontology about user's impairments (disability) related data, discussed in detail in [1]. The user's capability to carry out certain tasks is dependent upon the related impairment value in the impairments profile.

4.2 Memory Recall Pattern

A specialization of our proposed Generic Accessibility Pattern for helping the people with memory deficit is shown in Fig. 2. It shows the relationship between the user's ability to recall things, the lifetime information items which are associated with life events, and the representations to present or expose those events to the user.

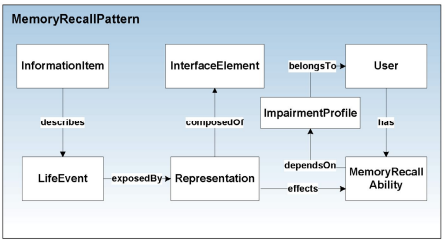


Fig. 2. Ontology Pattern for Inferring Effect on User's Memory Recall

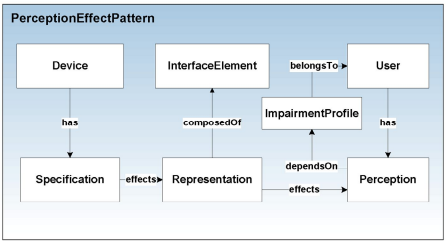


Fig. 3. Ontology Pattern for Inferring Effect on User's Perception

Following is the pattern's explanation in context of our *SemanticLIFE* system:

- *InformationItem*: Feed items such as emails, files of multiple types, web browsing history, chat sessions, processes running on user's PC, contacts, calendar.
- *LifeEvent*: Important events in user's lifetime which are helpful in recalling other entities and events such as birthdays, anniversaries, and other important occurrences. The events are identified and explicitly specified by semi-automatic analysis and annotation of the information items. It is important to keep in mind that there exists an *m:n* relationship between lifetime information items and events.
- *Representation*: The interface or visualization which is used to show these life events to the user. There can be different visualizations for different users because, for example, the preferred events sequence and events identification may vary according to user's impairments profile.
- *MemoryRecallAbility*: The ability of the user to recall the events or entities such as person and location with the help of representation. It can be measured against some specific tasks performed by the user designed heuristically.

4.3 Perception Effect Pattern

The *Memory Recall Pattern* is a specialization of the generic accessibility pattern where as the *Perception Effect Pattern* (Fig. 3) is its variation. It depicts *cause-effect* relationship between the user impairments, device profile and interface capabilities. The components modified or added for this purpose are described below:

- *Device*: Interaction device or parts thereof; like keyboard, display and cell phone.
- *Specification*: Semantic description of the device, known as the device profile [16].
- *Perception*: User takes time for making sense of the presented information, and then understanding how it is fulfilling his / her intended task. It can effectively be measured quantitatively and qualitatively when connected with task ontology.

4.4 Outcome of Ontology Design Patterns

These patterns are to be implemented in OWL-DL to enrich the connecting ontologies for the following domain models:

- User information space covering information items and life events
- Representation or Visualization

- User impairments
- Domain related tasks
- Device profile

The above analysis and outcomes helped us to define the accessibility framework.

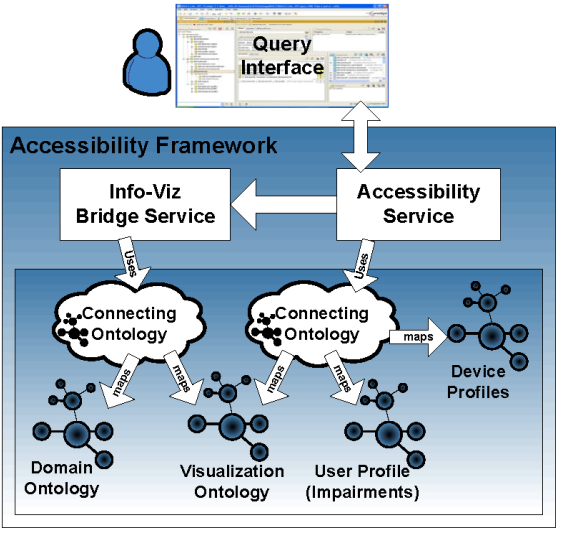


Fig. 4. Overview of Accessibility Framework

5 Accessibility Framework

The proposed framework (see Fig. 4) exploits *Connecting Ontologies* for its work. User's request is forwarded to the accessibility service that asks Info-Viz Bridge service to suggest, for example, appropriate interface elements or visualizations.

The recommendations of Info-Viz Bridge service are based on the rules and patterns established between visualization / representation and tasks ontologies. It is important to note that *we envision tasks as related to domain. Therefore those are modeled within the domain ontology.* The recommendations are sent back to the accessibility service. The interface is later adapted and made accessible using another connecting ontology which connects user impairments with device profiles and visualization recommendations through ontology design patterns and rules.

Our accessibility service may also be useful for providing accessibility extensions to existing ontology based UI frameworks such as Haystack [17]. Successive dry run of the inference system based on ontology design patterns and rules in connecting ontology were made to ensure its validity in real world scenarios.

6 Example Use Case

The connection between user interface characteristics and the user's visual impairments was explained in [1]. Here, the possibility to incorporate user's memory

recall pattern (see Fig. 2) in the *SemanticLIFE* system is shown. This is a relatively simple case of memory dysfunction problems caused due to old age or otherwise. The ontology patterns exercise resulted in following *Connecting Ontologies*:

- User information Items ↔ Life Events
- Life Events ↔ Representation
- User Impairments ↔ Representation
- Tasks ↔ Representation
- User Impairments ↔ Tasks

The main classes in ontology for user's lifetime information items which are being stored in our system are shown in Fig. 5.

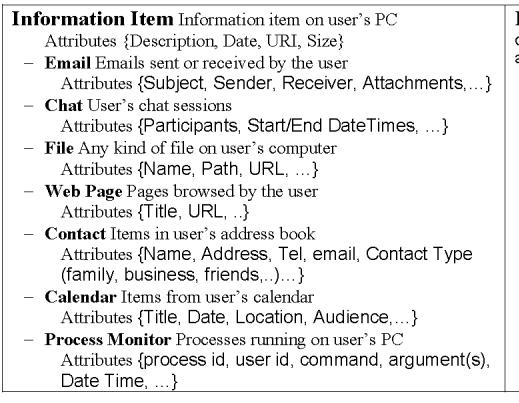


Fig. 5. Lifetime Information Items in Repository

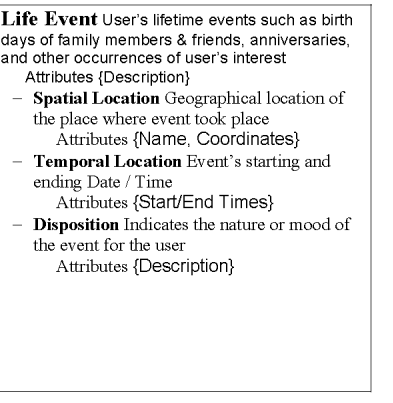


Fig. 6. Lifetime Events

The features of Life Events are described in Fig. 6. Life Events ontology is part of our comprehensive user model. The *Disposition* of an event can be categorized such as family, business, happy, sad, and important. *It is to be noted that our Life Events ontology will complement the events recorded in user's Calendar.* Some examples of the concepts which are part of the *Connecting Ontology* between lifetime events and lifetime information items are given below:

- ⟨InformationItem, *belongsTo*, LifeEvent⟩
- ⟨Contact, *participatesIn*, LifeEvent⟩
- Concept to specify role(s) of participants in the event
- Concept to identify the event of interest based upon user's preferences

In *Connecting Ontology* these concepts are formally specified using OWL-DL and the available rule languages are to be used for making sophisticated reasoning using Jess[7], and other reasoners such as Jena[8].

The state of knowledge about the domain of the *Connecting Ontology* being developed is very important. There are situations when the domain knowledge for the

[7] http://herzberg.ca.sandia.gov/jess/ (January 15, 2007)
[8] http://jena.sourceforge.net/inference/ (January 15, 2007)

CO exists either as (i) unstructured documents or as (ii) tacit knowledge with domain experts, or is (iii) still imaginative. In the use case under consideration, the domain knowledge about lifetime events exists as unstructured documents like emails, photo collections, file contents & chat session log. For generating *CO* the approaches under experimentation are based upon (a) text processing of *CO* domain knowledge and the two participating ontologies, or its refinement by (b) aligning participating ontologies with global standard ontologies, or (c) an approach exploiting all possible queries on the participating ontologies. However, the details are outside the scope of this paper.

7 Conclusions and Future Work

Bottom up approach of incorporating accessibility is a complex and unmanageable task which consumes a lot of resources. Consequently, it has proved to be a repulsive factor in itself for the software producers in providing universally accessible tools in general [18]. The presented approach tackles the problem space components' interactions from a holistic point of view using *Connecting Ontologies* while preserving the freedom of components' reuse by having individual ontologies of their own. Ontology design patterns are successfully being employed in our system which makes the approach more convincing for the practitioners.

This is a work in-progress for the *SemanticLIFE* project. Legal and ethical implications associated with a person's life time information, especially in case of memory and cognitive dysfunctions because of person's inability to make the right decision needs careful consideration and further research.

Acknowledgements. This work is supported jointly by the Higher Education Commission of Pakistan, the ASEAN-EU University Network Initiative and the Austrian Government's FIT-IT project DynamOnt under the contract FFG 809256/5512. We owe special thanks to Zubair Kareem, M.D. Phy-Neurologist (Holyoke, MA) for the useful feedback concerning cognitive impairments.

References

1. Karim, S., Tjoa, A.M.: Towards the use of ontologies for improving the user interaction for people with special needs. In: Miesenberger, K., Klaus, J., Zagler, W., Karshmer, A.I. (eds.) ICCHP 2006. LNCS, vol. 4061, pp. 77–84. Springer, Heidelberg (2006)
2. Chen, H., Colaert, D., Roo, J.D.: Towards adaptable clinical pathway using semantic web technology. In: W3C Workshop Semantic Web for Life Science, Position paper for W3C workshop Semantic Web for Life Science (2004)
3. Haller, A., Oren, E., Kotinurmi, P.: m3po: An ontology to relate choreographies to workflow models. In: Proceedings of the IEEE International Conference on Services Computing (SCC'06), Washington, DC, USA, IEEE Computer Society, pp. 19–27 (2006)
4. Lenzerini, M.: Data integration: a theoretical perspective. In: PODS '02: Proceedings of the 21st ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems, pp. 233–246. ACM Press, New York, USA (2002)

5. Sheth, A., Agrawal, S., Lathem, J., Oldham, N., Wingate, H., Yadav, P., Gallagher, K.: Active semantic electronic medical record. In: Cruz, I., Decker, S., Allemang, D., Preist, C., Schwabe, D., Mika, P., Uschold, M., Aroyo, L. (eds.) ISWC 2006. LNCS, vol. 4273, pp. 913–926. Springer, Heidelberg (2006)
6. Bechhofer, et al.: OWL Web Ontology Language Reference. W3C Recommendation http://www.w3.org/TR/2004/REC-owl-ref-20040210/
7. Seaborne, A.: RDQL - A Query Language for RDF, W3C Member Submission http://www.w3.org/Submission/2004/SUBM-RDQL-20040109/
8. Bouquet, P., Giunchiglia, F., van Harmelen, F., Serafini, L., Stuckenschmidt, H.: Contextualizing ontologies. Journal of Web Semantics Web Semantics: Science, Services and Agents on the World Wide Web 1(4), 325–343 (2004)
9. Gruber, T.: Toward principles for the design of ontologies used for knowledge sharing. Human-Computer Studies 43(5-6), 907–928 (1995)
10. Smith, M.K.,Welty, C., Mcguinness, D.L.: Owl web ontology language guide. W3C Recommendation http://www.w3.org/TR/2004/REC-owl-guide-20040210/
11. Kutz, O., Lutz, C., Wolter, F., Zakharyaschev, M.: E-connections of abstract description systems. Artificial Intelligence 156(1), 1–73 (2004)
12. Baader, F., Lutz, C., Sturm, H., Wolter, F.: Fusions of Description Logics and Abstract Description Systems. Journal of Artificial Intelligence Research (JAIR) 16, 1–58 (2002)
13. Kalfoglou, Y., Schorlemmer, M.: Ontology mapping: the state of the art. The Knowledge Engineering Review 18(1), 1–31 (2003)
14. Gangemi, A.: Ontology Design Patterns for Semantic Web Content. In: Gil, Y., Motta, E., Benjamins, V.R., Musen, M.A. (eds.) ISWC 2005. LNCS, vol. 3729, pp. 262–276. Springer, Heidelberg (2005)
15. Tjoa, et al.: SemanticLIFE - A Framework for Managing Information of a Human Lifetime. In: Proceedings of the 6th International Conference on Information Integration and Web-based Applications and Services (2004)
16. Klyne, G., Reynolds, F., Woodrow, C., Ohto, H., Hjelm, J., Butler, M.H., Tran, L.: Composite Capability/Preference Profiles: Structure and Vocabularies 1.0, W3C Recommendation http://www.w3.org/TR/2004/REC-CCPPstruct- vocab-20040115/
17. Karger, D.R., Bakshi, K., Huynh, D., Quan, D., Sinha, V.: Haystack: A Customizable General-Purpose Information Management Tool for End Users of Semistructured Data. In: CIDR (2003)
18. Keates, S.: Pragmatic research issues confronting HCI practitioners when designing for universal access. Universal Access in the Information Society 3, 269–278 (2006)

Immersive Digital Games: The Interfaces for Next-Generation E-Learning?

Michael D. Kickmeier-Rust[1], Neil Peirce[2], Owen Conlan[2], Daniel Schwarz[3], Dominique Verpoorten[4], and Dietrich Albert[1]

[1] Department of Psychology, University of Graz, Universitätsplatz 2 / III, 8010 Graz, Austria
{michael.kickmeier,dietrich.albert}@uni-graz.at
[2] Department of Computer Science, Trinity College Dublin, Ireland
{peircen,Owen.Conlan}@cs.tcd.ie
[3] Laboratory for Mixed Realities, Cologne, Germany
schwarz@lmr.khm.de
[4] Support Laboratory for Telematic Learning, University of Liège, Belgium
dominique.verpoorten@ulg.ac.be

Abstract. The intrinsic motivation to play, and therefore to learn, that might be provided by digital educational games teases researchers and developers. However, existing educational games often fail in their attempt to compete with commercial games and to provide successful learning. Often some learning is added to digital games or some gameplay is added to educational applications. Successful educational games, however, require merging professional game design with sound pedagogical strategies, creating a new hybrid format. Moreover, a methodology is required that allows continuously balancing gaming and learning challenges and the learner's abilities and knowledge in order to retain an immersive gaming experience. In this article we introduce approaches to game design and didactic design, as well as a framework for adaptive interventions in educational games.

Keywords: Digital educational games, Game design, Didactic design, Adaptive intervention, Microadaptivity, Competence-based Knowledge Space Theory.

1 Introduction

Over the past decades e-learning and technology enhanced learning and teaching, has become more and more popular. Consequently, a vast number of e-learning platforms and multimedia learning objects are available and widely accepted for the use in schools as well as for continuing education. E-learning has high potential to distribute high quality learning objects, facilitating learning and enabling new perspectives and insights to learning content.

However, the majority of e-learning and multimedia learning objects are based on traditional 2D user interfaces, e.g. website interfaces, Flash animations, Java applets, or PDF files. Provocatively speaking, current e-learning offers have all more or less the same unexciting look and feel. This perspective is compounded by the

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 647–656, 2007.
© Springer-Verlag Berlin Heidelberg 2007

proliferation of immersive computer games that learners are exposed to outside of learning experiences. Moreover, independent of usability and accessibility, traditional interfaces for educational applications have serious weaknesses from the perspectives of learning psychology and didactics. For example, they are not intrinsically motivational and it is difficult to retain a learner's interest, it is difficult to provide a meaningful context throughout learning episodes, it is difficult to activate prior knowledge as a basis for learning, it is not always possible to provide real-world problems for practicing new knowledge, and meaningful application of new knowledge is difficult without a meaningful and engaging context [1].

Immersive educational computer games offer a highly promising approach to overcome the mentioned weaknesses and to make learning more engaging, satisfying, and probably more effective. Currently, there is significant hype over game-based learning, ranging from edutainment to games for primarily educational purposes (see [2] for an overview). The major strength of digital games in education is a high level of intrinsic motivation to play and to proceed in the game and, thus, to learn within the context of a meaningful and continuous storyline and within the related para-scoial dimension provided by game characters. According to Malone [3], the factors forming that strength and making games fun are challenge, fantasy, and curiosity. Educational games provide clear goals and rules, a meaningful learning context, an engaging storyline, immediate feedback, a high level of interactivity, challenge and competition, random elements of surprise, and rich and appealing learning environments [3], [4]. These factors determine motivation to play and to learn but are also considered to be important for successful and effective learning, e.g. interactivity, feedback, problem solving, or context effects (for reviews see [1], [5]). On the other hand, educational games have also major disadvantages like difficulties in providing appropriate balance between gaming and learning activities, providing a continuous balance between challenge and ability, aligning the game with national curricula, or the extensive costs of developing high quality games (for a review see [6]). Due to these problems, most of today's educational games cannot compete with their commercial counterparts in terms of gaming experience, immersive and interactive environments and storytelling, or intrinsic motivation to play. Moreover, most educational games do not rely on sound instructional models thus leading to a separation of learning from gaming and often they provide gaming actions only as reward for learning. Therefore, such games do not differ significantly from other traditional multimedia learning applications.

The ELEKTRA project (www.elektra-project.org) has the ambitious and visionary goal to fully utilize the advantages of computer games and their design fundamentals for educational purposes and to address and eliminate the disadvantages of game-based learning as far as possible. Nine interdisciplinary European partners contribute to the development of a sound methodology for designing educational games and the development of a comprehensive game demonstrator based on a state-of-the-art 3D adventure game teaching physics according to national curricula. Furthermore, ELEKTRA will address important research questions concerning game design, didactic design, or adaptive interventions.

2 Game Design

It is the point of view of many game designers, educationists, and psychologists that the genre of digital games offers wide range of means of expressions to enhance the presentation of educational contents and to provide learners with an exciting and new interface to knowledge. In the disciplines of game design and development, there exists a long tradition and extensive experience in the creation of immersive narratives and story lines as well as in the design of appealing virtual environments. Design domains include game play design, story design, visual design, information design, sound design, interface design, and interaction design, and each is a fine art. Moreover, an appropriate usability of the game interface must be assured [7]. When aiming for integrating educational purposes into digital games and developing educational games with the main goal of conveying educational content, the designers are facing completely new challenges and demands. On the other hand educationists are confronted with the potential of a highly interactive, spatiotemporal game world that goes far beyond the interfaces of traditional e-learning applications. Educational games require the establishment of a new genre within which both worlds, recreational games and psycho-pedagogical aims and methods, are melded into a new hybrid format. From an educational perspective, digital games offer a real-time 3D virtual environment that serves as a space for learning. Such an environment is determined by *situations*. In digital educational games, there exists three types of situations: (a) learning situations (LeS) with the particular aim of teaching content, (b) gameplay situations that provide interactions with the game's environment, objects, and characters, requiring a skilful usage of tools and artfully accomplishing game moves, and (c) story situations that are required to combine learning and gameplay situations in a motivating and appealing way. The different situations and types of situations must be combined, driven, and justified by a global narrative that provides a meaningful context. A large number of open questions remain to be addressed by research within and beyond the ELEKTRA project. These questions include concerns over the appropriate balance of gameplay and learning, the translation of game elements into pedagogically sound learning methods, and the role of para-social interaction play (either with real mates or with non-player characters; NPC.)

3 Didactic Design

In addition to immersive and motivating game design, a sound pedagogical approach is required to develop a successful educational game. Choosing reference pedagogical frameworks that allow instructionally informed decisions is a key success factor when developing any kind of virtual learning environments [8]. The importance of such frameworks has been emphasized especially in the field of digital game-based learning [4], [8], [10]. Without such educational beacons, there is a high risk of a lack of a suitable vocabulary for describing and implementing pedagogical endeavors, thus ultimately loosing instructional focus. This section briefly presents the dominant conceptual tool used in ELEKTRA for dealing with methods of learning. The model is tentatively connected to Bloom's taxonomy of educational objectives [11]. The

articulation of methods, objectives, and evaluation is conceptually scrutinized according to the principle of "triple consistency".

Created in the context of teacher professional development ([12], [13]), the 8 Learning Events Model (8LEM) is intended to trigger practitioners to reflect on methods of learning and their diversification. This emphasis on methods is grounded in an educational philosophy which postulates that any educational action cannot restrict itself to "products of learning", but must also look at "processes of learning". 8LEM was conceived as a means for teachers to broaden their repertoire of methods. It helps in the systematic management of the difficult theoretical and practical issue of designing learning/gaming paths that are equivalent, be they different (inter-individual diversification) or complementary (intra-individual diversification [14]). The learning events, of which the paths are comprised, specified in 8LEM are exploration, imitation, reception, creation, experimentation, practice, debate, and meta-reflection.

Methods, adapted or not, do not take place in isolation. They become meaningful once they are connected to learning objectives [15]. Each 8LEM-labelled activity is attached to the levels of Bloom's taxonomy [11] and supposed to provide designers with a sound framework of instructional design [16]. The taxonomy includes learning objectives like knowledge, understanding, application, analysis, synthesis, and evaluation. The resulting combinations of learning events and objectives result in 48 learning activities. This gives a rough but complete overview of what is pedagogically possible in the virtual learning environment.

As mentioned above an important component of ELEKTRA's pedagogical conceptual tool-kit, supporting game development is the notion of "triple consistency". This cornerstone principle of valid pedagogy stresses the mandatory solidarity which must exist between objectives, methods and evaluation. Tyler [17] demonstrated that failures in learning system are related to a lack of consistency between learning objective, method used to teach, and the evaluation of how far the objectives have been met.

4 Adaptive Interventions

An immersive digital game for educational purposes requires a subtle balance of challenge and ability. Thus, the game must be able to adapt to the learner’s knowledge, skills, and abilities, motivation, and also pedagogical implications. Generally, adaptive approaches to e-learning contest the one-size-fits-all approach of traditional learning environments, trying to tailor the learning environment according to individual needs and preferences. Adaptivity refers to navigation, curriculum sequencing, and presentation. For example, an adaptive system may only provide learning objects which are suitable for an individual’s learning progress - learning objects either too difficult or too easy might not be displayed in order to avoid visual and cognitive load and to suggest an appropriate learning path through the learning content. In the context of immersive digital games, existing approaches to adaptivity must be extended in order to retain an immersive gaming experience, motivation, and probably flow experience by suitable adaptive interventions.

4.1 Adaptive Influencing of Learning Situations in a Game

Narratives and storylines have long been known to be a major motivating factor in video games. Contemporary games frequently employ a game narrative as a motivating factor. Examples can found through out all video game genres and include many popular game franchises such as *Tomb Raider, Command & Conquer, Grand Theft Auto, Super Mario,* and *Final Fantasy*[1] to name but a few. Evidently the presence of a plausible narrative within a video game can lead to an engrossing and rewarding experience. It is for this reason that the ELEKTRA demonstrator has an embedded narrative that has been crafted to appeal to the target audience of 13-15 year old school pupils. Whereas an embedded narrative has major advantages for motivating a learner to engage, it also poses considerable challenges for educational adaptation within the game. Educationally adaptive systems to date have focused largely on adaptively ordering and presenting learning tasks. Within an educationally adaptive game such as ELEKTRA the learning tasks are so integrated with the games narrative that the reordering of learning tasks would result is a corresponding reordering of narrative plot elements. With a linear narrative this would result in a nonsensical narrative that is implausible. The challenge of creating dynamic yet plausible adaptive narratives is considerable and requires arduous manual editing of branching narratives. Experimental systems such as *Façade* [18] exemplify the difficulties of creating adaptive narratives.

Within the field of adaptive hypermedia, adaptation is limited by the presentation medium and so adaptation is manifested through intermittent curriculum ordering and adaptive presentation (e.g., [19]). Due to the nature of 3D immersive games the adaptation within the ELEKTRA demonstrator needs to be continuous and less periodic; it needs to occur at a greater frequency than on a task by task level. Considering this with the existing difficulties associated with generating adaptive narratives, the ELEKTRA game provides microadaptivity, that is, adaptation within learning situations as opposed to around them.

Microadaptivity removes the challenges of adaptive narratives yet creates challenges of its own due to the nature of the experience of game play, and the impact that game world changes can have on a player's experience. Games are considered to be intrinsically rewarding to play, and it is factors such as self-governance, immersion, clear goals, immediate feedback, and game feasibility that contribute to rewarding play. Any adaptation within the game world must thus be achieved in such a manner so that it is not to the detriment of the play experience. Essentially the experience of *Flow* [20] must be maintained, where the learner is immersed, self-governing, aware of their ability to succeed, and clearly able to see the game's goal. From an educational perspective the experience of Flow is rarely catered for in adaptive eLearning systems. There is evidence however to suggest that Flow can be beneficial to learning in classroom environments [20] and may prove beneficial to learning experiences in immersive educational games.

[1] Tomb Raider, Core Deign (1996); Command and Conquer, Westwood Studios (1995); Grand Theft Auto, DMA Design (Rockstar North) (1997); Super Mario Bros., Nintendo (1983); Final Fantasy, Square Enix (1987).

4.2 Non Invasive Microadaptivity

Achieving microadaptivity without compromising the learner's gaming experience requires that any adaptation must be achieved in consideration of the gaming experience. To accomplish this within ELEKTRA the Learning Engine, the system responsible for adaptation, does not directly intervene with the game engine but provides recommendations for changes that are desirable. These recommendations are always contextually specific to a learning situation and are manifested in such a way that they are in line with the game's narrative and the learner's sense of self-governance. The intention of the Learning Engine as to what needs to be adapted is embodied by the recommendations. However, it is at the ultimate discretion of the game engine as to whether or not to enact a recommendation, this is provided as an extra insurance so that no adaptation is detrimental to the learner's gaming experience.

It is through these appropriate recommendations that the Learning Engine can effectively influence the game engine to adapt without negatively impacting the game's narrative. The ultimate realization of the recommendations is determined by the game engine and they can take the form of hinting or guidance from a NPC, motivational support from an NPC, throttling the learning situation difficulty, or appropriate direct intervention by an NPC where deemed necessary.

4.3 Abstraction Within the Learning Engine

The ELEKTRA game consists of a number of LeS each of which may be adapted to the learner's performance. The generation of adaptive rules for each learning situation independently would be both time consuming and lead to many rules that are similar in nature yet sufficiently LeS specific to prevent their reuse. This is undesirable from an authoring perspective as it is both arduous and difficult to author adaptation for each learning situation; additionally the reuse of adaptive pedagogical strategies is hindered. The ELEKTRA demonstrator overcomes this through a system of abstracted pedagogical adaptation. Whereas with many adaptive educational systems the adaptation for a learning situation is specific to that situation only, the Learning Engine provides generic abstracted pedagogical adaptation that is applied to all learning situations. This is then mapped into recommendations that are specific to a learning situation. It is in this mapping from generic adaptation to specific game changes that allows the Learning Engine adaptively influence a learning situation without adversely affecting the game's narrative.

The learning engine's adaptation methodology can be broken into four separate components, broadly described as Evidence, Inferring, Recommendation, and Realization. The first component, Evidence, is that information coming from the game engine that is deemed pertinent to the learner's performance within a learning situation. The second component, Inferring, embodies an assessment of the Evidence information and the generation from this of generic semantically meaningful performance information. The Recommendation component consumes the output of the Inferring component in order to generate a generic pedagogical adaptation goal. The final component the Realization consumes generic pedagogical goals and produces learning situation specific adaptation within the game engine to achieve the

generic goal. The advantages of this specific-generic-specific cycle can be found in the ability of the system to reuse pedagogical rules, reduce the specific per learning situation adaptation authoring, and aid the independent authoring of both pedagogical adaptation and game learning situations.

In ELEKTRA the above cycle uses skills assessment at the Inferring component as the key means for moving from the specifics of a LeS to the generic capability to reason pedagogically about the learner's performance. For example, based on the learner's interactions with a learning situation evidence is produced for how they are progressing. This evidence is translated into their acquisition of skills pertinent to the domain in which the LeS is set. With no knowledge of the specific LeS, recommendations are made based on this inferred progress. These recommendations are pedagogically informed and indicate the type and level of influence required. It is then the responsibility of the Game Engine to determine how this recommendation will be realized.

4.4 The Cognitive Framework Behind Adaptivity

The cognitive framework underlying adaptivity in ELEKTRA is CbKST. In its original formulation, *Knowledge Space Theory* (KST) [22] [23] provides a set-theoretic framework for organizing a domain of knowledge and for representing the knowledge of individuals. The basic idea behind KST is to assume *prerequisite relations* among a finite set of problems (e.g., test-items). If a problem *a* is a prerequisite for a problem *b*, from mastering problem *b* also the mastering of problem *a* can be inferred. Based on the set of problems and the corresponding prerequisite relation a *knowledge structure* is established which includes all admissible or meaningful *knowledge states*. Each individual can be assigned to a specific knowledge state. The behavioral KST was extended by explicitly referencing the latent skills and competencies underlying the observable behavior [24], [25], [26]. One of the most successful approaches to CbKST is the *Competence Performance Approach* (CPA) by Klaus Korossy [27]. The basic idea of CPA is to assume a basic set E of abstract cognitive skills that are relevant for mastering the problems of a domain. The *skill state* of an individual is the collection of all available skills, which is not directly observable but can be uncovered on the basis of the observable performance on the problems representing the domain. As in KST, prerequisite relations are defined on the set of skills establishing a *skill structure C*, which contains all possible skill states. Utilizing *skill* and *problem functions*, families of subsets of skills can be mapped to problems, which can be mastered with the given set of skills and vice versa. By the assignment of skills to the problems of a domain, also a *performance structure* or *problem structure* on the set of problems is induced.

An example is given in Fig. 1. If we assume that skill *s1* is a prerequisite for skill *s2*, skill states, which include skill *s2* but not all of its prerequisites, in this example *s1*, are not admissible, at least from a deterministic perspective. This procedure allows us to substantially reduce the number of possible skills states. Without prerequisite relation, six skills establish a skill structure that includes 64 skill states, i.e. the power set of all skills (2^6). The prerequisite relation in Fig. 1 establishes a skill structure that includes only eleven possible skill states. A major advantage is that CPA allows multiple, individual, not necessarily linear learning paths. This offers a sound model

of learning objectives which, in turn, establishes the very basis for game design as well as learning design. Moreover, by an assignment of latent skills to LeS and single objects (e.g., tools, switches, laboratory equipment, or books) within LeS using problem and skill functions, the skill structure serves a basis for reasoning about the skill state of a learner on the basis of the performance in LeS.

Adaptively providing a learner with suitable learning objects is a major mechanism of existing approaches to adaptive e-Learning. This principle can be utilized for providing adaptivity on a macro-level in digital educational games. In the context of game design, branched storylines for interactive, personalized storytelling could be realized, characterized by multiple paths through the game environment. As discussed in section 4.1, in the context of educational games the principle of macroadaptivity has major drawbacks, however. First, it might be highly expensive to implement the requirements of adaptive storytelling, which is a demand of an adaptive ordering of LeS. Second, a plausible re-ordering might not be possible in a linear storyline and third, a theoretical model for adaptive storytelling still has to be developed. For these reasons, in the ELEKTRA project we pursue an approach to adaptivity on a micro-level.

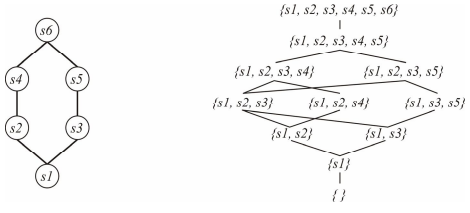


Fig. 1. Prerequisite relation (left) and the corresponding skill structure (right)

In order to meet the requirements of non invasive adaptive interventions, traditional CbKST-based approaches to the assessment of skills and competencies must be complemented with an approach of probabilistic competence diagnosis. In contrast to traditional approaches to adaptive knowledge assessment, which is mostly based on typical test items (e.g., a multiple choice task), adapting to a learner's skills on a micro-level requires a probabilistic assessment of knowledge based on the learner's actions within a LeS. Such LeS include a number of objects, which can be manipulated by the learner in order to accomplish a more or less complex problem solution task. For example, a task might be to focus a cone of light using available objects such as an electric torch and a set of convex and concave lenses. An efficient, probabilistic assessment of skills in this context requires interpreting the learner's actions and, thus, the sequence of positions (and alignments) of all objects. The positions of all objects of a LeS constitute a *problem solution state* that can be evaluated with a specific utility function, determining the correctness of the current problem solution state, or in other terms, the learner's approach to the task's solution. In order to assess the available and lacking skills of the learner, the skills, which are necessary to adequately manipulate an object, must be assigned to each object. The probabilistic skill assessment occurs by updating an initial probability distribution over all skill states based on the utility value of the evaluation of the problem solution

state. For example, if the learner uses the correct lens in order to focus a cone of light (i.e., a convex lens), the probabilities for skill states including the skill of knowing that convex lenses focus light, are increased. This procedure enables detecting the available and lacking skills of a learner. Moreover, on this basis also misconceptions (e.g., using a concave lens to focus light) or incorrect loops in the problem solution process (e.g., if the learners tries the same incorrect manipulations over and over again) can be detected. On this basis the system can determine suitable interventions (e.g., hints) for a specific problem solution state or sequence, which are directly related to the underlying latent skills.

5 Conclusion

The ELEKTRA project has the visionary goal of making a significant step towards educational games that can compete with commercial games in terms of immersion, engagement, and motivation while being a full educational option, utilizing appropriate balance between learning and gaming and relying on national curricula and sound psychological, pedagogical, and instructional theories. Developing a successful educational game, especially for older children, cannot be accomplished by adding some extra learning to a recreational game and it cannot be accomplished by adding some gaming to a traditional learning object. This aim requires the generation of a new *genre* within which both worlds, recreational games and psycho-pedagogical aims and methods, are melded. The adaptive technology which already exists and which is extended will help to fully incorporate the advantages of digital games for educational purposes and to make a significant step towards developing the interfaces for next-generation e-learning and technology enhanced learning.

Acknowledgments

The research and development introduced in this work is funded by the European Commission under the sixth framework programme in the IST research priority, contract number 027986.

References

1. Merrill, M.D.: First principles of instruction. Educational Technology Research and Development 50, 43–59 (2002)
2. Mitchell, A., Savill-Smith, C.: The use of computer and video games for learning: A review of the literature. Learning and Skills Development Agency, London (2004)
3. Malone, T.W.: Toward a theory of intrinsically motivating instruction. Cognitive Science 4, 333–369 (1981)
4. Prensky, M.: Digital game-based learning. McGraw-Hill, New York (2001)
5. Schulmeister, R.: Didaktisches Design aus hochschuldidaktischer Sicht - Ein Plädoyer für offene Lernsituationen. In: Rinn, U., Meister, D.M. (eds.) Didaktik und Neue Medien. Konzepte und Anwendungen in der Hochschule. Medien in der Wissenschaft vol. 21, pp. 19–49 (2004)

6. Van Eck, R.: Digital game-based learning: It's not just the digital natives who are restless. Educause Review 41, 16–30 (2006)
7. Holzinger, A.: Usability engineering for software developers. Communications of the ACM 48, 71–74 (2005)
8. Koper, R., Olivier, B.: Representing the learning design of units of learning. Educational Technology & Society 7, 97–111 (2004)
9. Squire, K.: Cultural framing of computer/video games. Game Studies. The International Journal of Computer Game Research vol. 1 (2002)
10. Gee, J.P.: What video games have to teach us about learning and literacy. Palgrave Macmillan, New York (2003)
11. Bloom, S.: Taxonomy of educational objectives, the classification of educational goals – Handbook I: Cognitive Domain. McKay, New York (1956)
12. Leclercq, D., Poumay, M.: The 8 learning events Model. Release 2005. (Retrieved online February 5, 2007) from http://www.labset.net/media/prod/8LEM.pdf
13. Verpoorten, D.: The 8 learning events model: a pedagogic conceptual tool supporting diversification of learning methods. In: Proceedings of the 6th IEEE International Conference on Advanced Learning Technologies - (ICALT 2006) - Kerkrade, The Netherlands (2006)
14. Clark, J.M., Paivio, A.: Dual coding theory and education. Educational Psychology Review 3, 149–170 (1991)
15. Crawford, C.: The art of computer game design. McGraw-Hill, Berkley (1984)
16. Ip, A., Morrison, I.: Learning objects in different pedagogical paradigms. In: Kennedy, G., Keppell, M., McNaught, C., Petrovic, T. (eds.) Meeting at the crossroads: Proceedings 18th ASCILITE Conference, pp. 289–298. The University of Melbourne, Melbourne (2001)
17. Tyler, R.W.: Basic principles of curriculum and instruction. University of Chicago Press, Chicago (1950)
18. Mateas, M., Stern, A.: Façade, an artificial intelligence-based art/research experiment in electronic narrative (Retrieved February 5, 2007) from http://www.interactivestory.net/
19. DeBra, P., Calvi, C.: AHA! An open Adaptive Hypermedia Architecture. The New Review of Hypermedia and Multimedia 4, 115–139 (1998)
20. Csikszentmihalyi, M.: Flow: The psychology of optimal experience. Harper and Row, New York (1990)
21. Shernoff, D.J, et al.: Student engagement in high school classrooms from the perspective of flow theory. School Psychology Quarterly 18, 158–176 (2003)
22. Doignon, J.-P., Falmagne, J.-C.: Knowledge spaces. Springer-Verlag, Berlin Heidelberg New York (1999)
23. Albert, D., Lukas, J.: Knowledge spaces: Theories, empirical research, and applications. Lawrence Erlbaum, Mahwah (1999)
24. Albert, D., Held, T.: Component based knowledge spaces in problem solving and inductive reasoning. In: Albert, D., Luks, J. (eds.) Knowledge Spaces: Theories, Empirical Research Applications, pp. 15–40. Lawrence Erlbaum, Mahwah (1999)
25. Doignon, J.-P.: Knowledge spaces and skill assignments. In: Fischer, G., Laming, D. (eds.) Contributions to Mathematical Psychology, Psychometrics and Methodology, pp. 111–121. Springer-Verlag, Heidelberg (1994)
26. Düntsch, I., Gediga, G.: Skills and knowledge structures. British Journal of Mathematical and Statistical Psychology 48, 9–27 (1995)
27. Korossy, K.: Modelling knowledge as competence and performance. In: Albert, D., Lukas, J. (eds.) Knowledge Spaces: Theories, Empirical Research Applications, pp. 103–132. Lawrence Erlbaum, Mahwah, NJ (1999)

Methodology and Application in a Technologically Enhanced Environment: The Case of the Writing Center at the Hellenic American Union and the Hellenic American University

Vassiliki Kourbani[1], Yiannis Petropoulos[1], and Dimitris Tolias[2]

[1] Hellenic American University, 12, Kaplanon str. 10680, Athens, Greece
{vkourbani,jpetrpop}@hau.gr
[2] Hellenic American Union, 22, Massalias str. 10680, Athens, Greece
dtolias@hau.gr

Abstract. The newly established Writing Center at the Hellenic American Union and the Hellenic American University is a breakthrough in what conventional Writing Centers entail in terms of both target audience and nature. It aims to address the needs of both the Hellenic American Union and the Hellenic American University students along with the general public for the improvement of academic, professional and writing The traditional role of Writing Centers-both in the USA and in Europe- is to offer academic writing assistance to students and help them develop their writing skills, for specific purposes or across the curriculum. The Writing Center at the Hellenic American University shares the same fundamental aims but extends its mission to a wider scope of applications and audiences, addressing the needs not only of the Hellenic American University students but also of the general public. Its services go beyond the traditional face-to-face, "paper-pencil" approach since assistance is provided through innovative design, e-equipment, and original software that encourage collaboration and individualized monitoring, both on site and on line. Thus, the Writing Center at the Hellenic American University aspires to become a showcase of how collaborative writing – defined in a variety of ways – is realized in terms of the cooperation among the particular Writing Center stakeholders and the learning outcomes involved.

Keywords: Writing Center, Distance Tutoring, Tutoring, Classroom Learning.

1 The Writing Center at the Hellenic American Union/University

A Writing Center has traditionally meant a service provided, usually at the university level, where individuals may come in and get one-on-one help on their writing. The Writing Center at the Hellenic American Union/Hellenic American University is a breakthrough in what a traditional Writing Center entails in terms of both target audience and nature. As an innovative educational and vocational service, the Writing Center combines cutting-edge technology and design with experienced personalized tutoring, making it unique in Greece and one of the few such services in the world.

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 657–666, 2007.
© Springer-Verlag Berlin Heidelberg 2007

Its mission is to provide support to Hellenic American Union and Hellenic American University students, in addition to outside patrons, on a one-to-one basis, for any type of writing problem at any point of the writing process.

The Writing Center endeavors to make the development of writing skills an important part of the writing experience, increase student literacy, and promote intellectual growth. Offered also as a free service to the University's students, it has been designed to cater to academic and vocational needs and address any type of writing problem. This is achieved, on site and on line, by combining innovative design, e-equipment, and cutting-edge software with experienced personal tutoring. Users have the option to access software that automatically corrects and comments on their submitted work online, round the clock, without having to be in the Center's premises.

1.1 Rationale

According to Hendry & Harper (1997), ''design is a complex activity in which a design for an 'artefact' emerges through a series of interactions between designers, information germane to the design and the design itself'' (p.50). A very important parameter for successful collaborative design and learning is ''shared access to, and management of, information needed to support design and learning'' (Hendry & Harper, 1997, p.50). Actually, nowadays as services become more complex, collaborative design is ''becoming more multidisciplinary and conducted by large team of specialists'' (Ferraro, Rogers & Geisler, 1995, p.120). Models of the collaborative learning process have increased rapidly in numbers focusing on creating a common knowledge base without the use of electronic tools. However, more recently collaborative sessions incorporate cutting-edge technology and the relevant software as a powerful environment for learning (Coulson & Feltovich, 1996).

Since collaboration is a crucial part of professional and personal life, institutions of higher learning prepare students for collaborative work. Actually, in a collaborative design session without switching screen, students are able to talk to each other, to draft common responses to questions, to work through differences of opinion, to send material, to share material with other users, to chat. Thus, by emphasizing student involvement and group participation and accommodating their own pace and learning style, it provides a mechanism to ensure flexibility, relaxation and practicality.

1.2 The Design of the Writing Center

Based on the principle of Collaborative Design, the Writing Center area, patented furniture, and equipment have been organized to allow for students' electronic cooperation in pairs or groups, without the prohibiting traditional one-on-one tutoring. The Writing Center area aims at ''accommodating the needs of intellectual teamwork by providing teams of users with the technological support required to generate , coordinate and refine the joint action required by common projects'' (Ferraro, Rogers & Geisler, 1995, p.121). This type of design requires close cooperation among the members of a team who take part in this unique learning environment. Actually, students come to the Writing Center -as a team- to share their ideas, put them into practice, and to understand material or complete a task (Sherry & Myers, 1998).

2 Learning at the Writing Center

2.1 The WRC as a Distance Tutoring Environment

Goodman (1984) in his theory of learning argues that what we call the world is really our interpretation of experience. As new experiences seem to cause us to change our perceptions, we tend to develop new symbols for expressing these understandings. Invariably, explorations of learning raise questions about ways of knowing, and these, in turn, lead to discussions concerning the qualities of intelligence. These discussions, which could be argued, are particularly critical in attempts to develop learner-centered communities because, to be learner-centered, learning communities must attend to individual difference, whether experiential or innate.

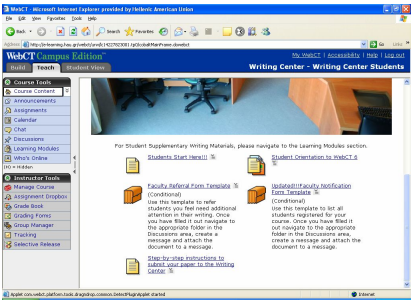
Collaborative theories of learning tend to view learning as a socially interactive process that is enriched by interpersonal communication. This construct therefore emphasizes the importance of knowledge that students bring to their new experiences and the value of shared perspectives which, in effect, draw substance from different ways of knowing (Mezirow, 1991). Plater (1998) suggests that communication technology is the most expedient vehicle for change since technology embraces distance learning to allow flexibility in meeting students' time constraints. To this respect, it is suggested that time could be viewed as a resource to allocate according to community objectives, emphasizing students learning outcomes and activities. The Writing Center has created a space in WebCT to offer an online component for tutoring. This space is provided for the faculty and the students of the University to communicate with and submit assignments to WRC tutors.

Methodology. The procedure followed by Hellenic American University instructors, students, and WRC tutors in relation to distance tutoring through the WebCT is as follows:

Instructors wishing to refer students to the Writing Center would log in to the WRC WebCT page with their username and password and click on the link to download the **Referral Form** template. Subsequently, instructors fill in the referral form with the name(s) of the student(s) they wish to refer to the WRC, along with a brief description of the student assignment they would like WRC tutors to deal with. Clicking on *Discussions,* a list of names will appear and within each name the current academic term, ex. *Fall 2006 Referral Forms*. Instructors then click on the link and upload the referral form submitting it to the Writing Center Staff. Afterward, instructors can log back in to review feedback sent to them by the WRC clicking on *Discussions*, locating the message posted by WRC staff, opening it and downloading the attachment to their computer.

Students referred to the WRC have a week's time frame to contact the WRC through the Web CT electronic platform to submit their work. Students can log in to the Web CT homepage using their username and password. Then they can locate the course for which they would like to submit their assignment, click on the link, and attach their assignment document. Students have the option to add comments on their assignment upon submission, if they feel they need specific help on a particular area.

Tutors receive submitted work through the e-platform and respond to both students and instructors. Once WRC tutors have worked on instructors' and students' concerns



relating to the submitted assignment they re-submit the reviewed document to students. At the same time WRC tutors re-submit the referral form to course instructors confirming the virtual visit, providing the date, type of assignment, and reference to main points reviewed. The WRC collects data on University students concerning frequency of visits, specific problem assessment, and overall writing performance. Students log back in to check whether they have received feedback from the Writing Center. In case feedback has been sent to the students' WebCT homepage, there will be two files attached, the first being the original assignment attached and sent to the WRC, the second being the reviewed assignment document re-submitted by the WRC tutors.

2.2 The WRC as an Onsite Tutoring Environment

Following Angelo's (1997) concept of learning, the writing center's virtual and distance component also offers both faculty and students the opportunity to learn from and teach one another enabling and promoting teamwork and interpersonal involvement and sustaining the "cooperative academic culture" that the Hellenic American University advocates.

When e-learning first became widespread, one of its biggest failures was the assumption that it could become a solution to all learning problems: that faculty was no longer required, and that anything could be e-taught. Success was only about broadcasting good quality learning material. Now we know that this is no longer the case, and that good teachers, communication, collaboration and exploration are essential. The challenge for educators is how best to use the available tools in order to emphasize personalization and effectiveness. The Writing Center tutors have a wide range of resources at their disposal as well as a clear understanding of their learners' needs.

Methodology. Course instructors send a referral form to the Writing Center through the WebCT e-platform. Then they provide a copy of that referral form to the student to be brought to the WRC for the on- site session. Instructors receive feedback from the WRC tutors after the completion of the on-site session through WebCT in the same process as in the case of distance tutoring sessions.

Students who wish to have an on-site tutoring session are given a week's time frame to make an appointment and visit the WRC. The appointment request is made through an e-mail message sent by students to the WRC electronic mailbox, and the WRC response to this request is sent again by e-mail to the students.

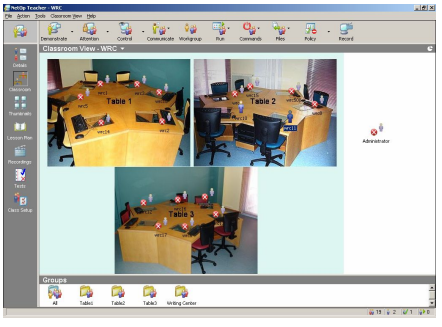
On the day of the on site session students visit the Writing Center premises with a copy of the referral form completed by their instructor(s). The session takes place making absolute use of all software applications available at the WRC learning and tutoring environment.

2.3 Application: WAC and the WRC

Information technology seems to be continuously important in education, not just as a subject in its own right, but also as a key enabler in improving the teaching of a variety of subjects and academic disciplines (Philpott, 2006, p.42). Furthermore, Chickering et al (1987) argue that technology may further promote the development of a learner-centered community. Under this framework, technology encourages contact and interaction between students and faculty, promotes cooperation among students, facilitates active learning through real-time and constructive feedback respecting simultaneously the diverse talents and learning styles of each individual student in any classroom environment.

The role of the Writing Center is central to the application of WAC in the Undergraduate Program of the Hellenic American University. The WRC offers assistance with students' writing and helps monitor their progress in their writing competencies.

Methodology. All students registered for general education courses are required to make a set number of visits/on line submissions to the WRC in connection to the preparation of a course writing component that carries a certain weight towards the course final grade. At least one week before the time period that the instructor would like students to visit the WRC, the instructor will complete the Instructor Notification Form and upload it to the WebCT Online Writing Center Home Page. The Instructor Notification Form contains information concerning the dates of the time period the students must visit the WRC, a brief description of the assignment and the due date, and a list of students in the class. Students submit their work through WebCT. In case students wish to visit the WRC on-site after assignment submission, they have to make an appointment. On the due date of the assignment, the WRC will notify the instructor which students visited the WRC and which failed to do so. At the end of the semester, the WRC will provide feedback to the instructor in connection to the visits.

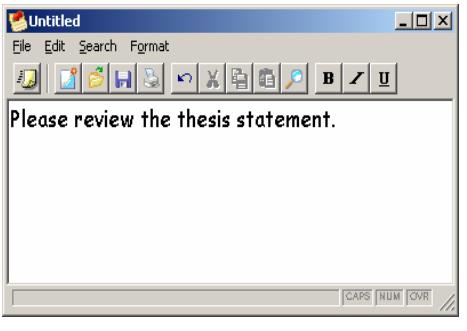


3 The WRC as a Classroom Environment

The Writing Center area is a specially designed room (*ibid. Design and Rational Section in this paper*). It consists of three '*learning and cooperation isles*'. Each isle accommodates six users facilitating interaction and collaboration interchangeably either as a group or in pairs. Each isle is octagonal shaped with all screens embedded in the conference table to allow for an ample work surface so as not to impede eye contact and communication among the teams. The Writing Center is equipped with eighteen (18) state-of-the-art student workstations, and one (1) instructor computer, a

Soft Board interactive electronic whiteboard, screen and projector. In the case of the HAU Writing Center, users have access to enhanced, online, instructive, visualized help in lieu of traditional writing centres. Students log into the system and with the help of the software programs available, they are connected with the instructor's computer to create an interactive, networked classroom.

Methodology. Writing instructors at the Hellenic American University also use the writing center as classroom environment in contrast to most writing centers which are primarily used for individual tutoring between instructors and students. In an effort to foster highly interactive learning experiences and build a shared knowledge base (Spiro, Coulson & Feltovich, 1996), five kinds of *processes* requiring intellectual teamwork are promoted. The first process is that of monitoring. It is important for the writing instructor to follow students' progress as they work at the Writing Center, and provide any feedback and assistance he/she deems appropriate and necessary. Using the NetOp School© software, the instructor can demonstrate, monitor, and interact with his/her students simply and effectively. This can be done on a *whole-class-basis*, to *selected groups*, or to *individual students*. The software allows the instructor to individualize his/her teaching meeting the needs of each one of the students in the classroom.



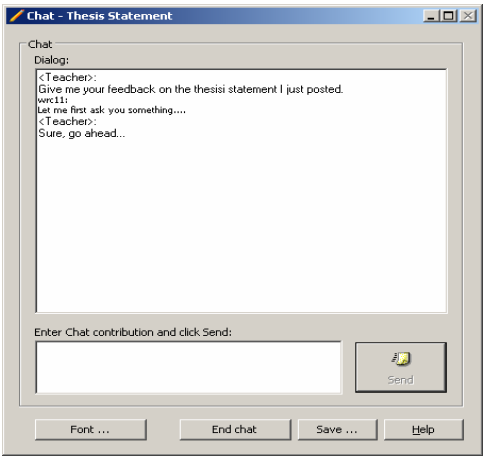
Since this is *a real-time* class monitoring system, the instructor simply selects from his/her screen the student computer that he/she wishes to monitor without interrupting the student's flow of work. While in monitoring mode, the instructor may help a student who faces a problem with his/her paper by sending him/her a written message (note) with advice and tips. The note reaches the recipient and appears on his/her computer screen. Each time, the instructor has the option to either select one student or a group of students as recipients of his/her notes. The student cannot reply to this note but he/she may copy the information on his/her computer for reference during his/her writing task. The message on the note is usually brief so that the student does not have to spend much of his/her time to deal with it. In most of the cases, the note includes a tip, pinpoints to an error in the text, or provides a web link to an external source for extra help.

The second process involves interaction[1]. Research (Sherry& Myers, 1998, LaMonica, 2001) shows that students become better writers when they have the option to interact and communicate with their fellow students and their instructor

[1] Current literature does not provide one standard definition of interactivity. Among others, Borsook and Higginbotham-Wheat (in: Sims, 1998) propose a set of interactivity elements which would not only classify a program as interactive but also help the learner: responses should be immediate; access to information should be non-sequential; programs should be adaptable, provide feedback and options to the end-user; programs should allow for bi-directional communication and should be characterized by a short presentation sequence, thus enhancing interactivity.

while drafting their compositions (Bull, Kimball& Stansberry, 1998). Interaction typically takes place in class discussions, small groups, student-teacher conferences (Verneil& Berge, 2000), and through written comments to the student. The Web, coupled with carefully selected computer software, provides new ways to the writing instructor to interact with students and enhance student-to-student communication. The Writing Center learning environment offers the instructor the opportunity to communicate with the students on every detail of the activity assigned to them during class time.

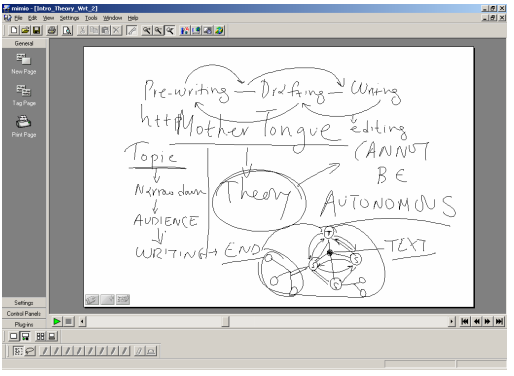
The NetOp, and the NetMeeting© software allow room for continuous interaction between the instructor and the student(s) in the form of written text, sound files, and images. Each student is able, at any given time, to access immediate feedback on the progress of his work, share it with other students or form subgroups within the same class following the instructor's advice and suggestions. The ability to interact and exchange ideas, and feedback accelerates learning and facilitates communication and cooperation among the students (Moore& Kearsley, 1996).



As a result, the instructor is in the position to form active working groups that take into consideration each student's needs and skills enabling organized learning with emphasis on individual attention and performance. The instructor may for instance use the Chat Feature of NetOp to communicate with one or as many of his/her students providing them with the option to reply to him/her and/or initiate a real time discussion with their classmates. The Chat and the Notes features remain active throughout the session time to enhance the exchange of information, ideas and feedback in the classroom.

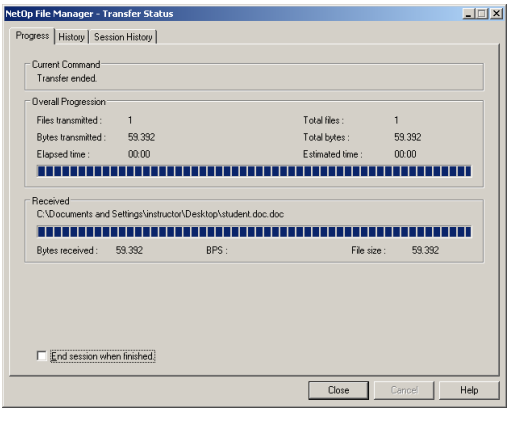
The third and fourth processes in the Writing Center methodology involve **collaboration** and **visualization**[2] . In collaborative learning, students work together on teams to maximize learning (Walker, 1997; Slavin, 1990) and the teacher acts as facilitator, coach and coordinator. Among the attributes of teamwork (Thomas&Funaro, 1990; Blaye et al., 1990) are a clear, positive interdependence among students, regular group self-evaluation, behaviors promoting each member's success and learning, individual accountability and responsibility and successful interpersonal and social skills. Encouraging students to work collaboratively in pairs or small groups in an effort to build upon and enhance knowledge, constitutes an important strategy in the teaching philosophy of the Writing Center.

[2] Images capture the thrills, emotions, and concerns of people. Most people depend on ''visual input for much of their understanding of the world around them and as a basis for further creative and more complex activities'' (Shneiderman, 2003, p.90). The visualization of the information allows designers to succeed in showing and – in essence - "telling" hundreds of times more information than tabular displays''(Shneiderman, 2003, p.90).



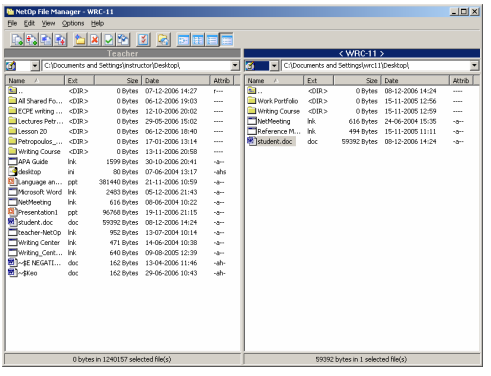
One of the first steps towards the materialization of this philosophy is the establishment of a constructivist environment (Lapadat, 2002; Whittle, Morgan, & Maltby, 2000) wherein students create projects, building their own knowledge. The process of selecting teams and having students work together starts at the beginning of each course designed for the Writing Center. The Writing Center makes collaboration easier both for the instructor and the students. The use of the Mimio© software for example, which remains active from the beginning of each session, allows the instructor to have the complete attention of the class during lecturing since all teacher notes that are written on the board are projected to the central screen and are also saved and transmitted electronically onto the students' computers at the end of the class. During the lesson, students do not have to worry so much about whether they have received the appropriate notes from the instructor's lecture. Instead, they can concentrate on what the instructor says and writes, thus enhancing their active participation and understanding of the lesson.

The NetOp© software also facilitates immediate visualization and sharing of the work produced during class time by the students. The instructor can select particular members in the work groups of his/her class and have them share (with or without the ability to intervene) a text or a folder, assign an activity and provide feedback at every stage of the students' work. In other words, the instructor and the student can work on a document at the same time making corrections on the spot, suggesting alternatives and requesting for clarification and assistance.



The **transfer** of new knowledge quickly and effectively among the students is the fifth process applied in the Writing Center. It could be argued that one way to speed up the learning process is to use multiple channels effectively (Gilpin, 2000). Actually, the integration of multiple channels accelerates information distribution, expands educational options, and generally improves the quality of learning outcomes. The combination of complex interactions among multiple channels (i.e. software applications) promotes and enhances learning transfer. The Writing Center instructor mixes a variety of technology enhanced tools and software (see NetOp and Mimio) and other forms of appropriate input (i.e. audio and written text, video, etc) while teaching his/her class

and significantly speeds up the transfer of knowledge tailored both to the needs of the class and the individual student.



Using the NetOp or NetMeeting, for instance, the instructor can select and send text and or other media files to the students, either as a group or individually, selecting the appropriate means to this purpose. In the Writing Center learning environment, students are bombarded with multiple inputs meeting various learning styles. After all, it is the combined use of the interactive board, with the PowerPoint presentation, the monitoring of the screen, the real time communication and selected dispatching of information, and the option of publicizing a student's work in order to provide an example or sample to work on for the rest of the class, that actually transform this classroom into an knowledge and communication exploration environment (Richardson & Ting, 1999).

The Writing Center's learning objective is to place different pieces of knowledge in a single package to secure that this knowledge can be quickly and effectively transferred from one student to another within the environment. This new approach to the role of the Writing Center may require, for the instructor mainly, some time to manage and combine successfully and productively the various software and applications, but in the end it streamlines the knowledge transfer process; the most vital and integral component of every educational institution.

References

1. Angelo, T.A.: The campus as learning community. American Association for Higher Education, vol. 49(9) (1997)
2. Bull, K.S., Kimball, S.l., Stansberry, S.: Developing interaction in computer-mediated learning (ERIC Document Reproduction Service No.ED 417902) (1998)
3. Blaye, et al.: Collaboration as a facilitator of planning and problem solving on a computer-based task. CITE Report 90, Institute of Educational Technology, Open University, Milton Keynes, U.K. (1990)
4. Chickering, A.W., Gamson, Z.F.: Seven principles for good practice in undergraduate education. The Wingspread Journal, Special Section vol. 9(2) (1987)
5. Ferraro, A., Geisler, C., Rogers, E.H.: Team learning through computer supported collaborative design Computer Supported Collaborative Learning Conf (1995)
6. Hendry, D.G., Harper, D.J.: An informal information-seeking environment. Journal of American Society for Information Science 48(11), 1036–1048 (1997)
7. Gilpin, V.: Best practice in innovative learning: Information technology as a catalyst for change: A summary of recent Africa experience. Focus On Africa (2000) Retrieved February 27, 2003 from http://www.crite.net/conferenes/2000/iscl/Focus on Africa.html
8. Goodman, N.: Of mind and other matters. Harvard University, Cambridge MA (1984)

9. LaMonica, L.: The role of the instructor in web-based instruction: Are we practicing what we teach? (2001) Retrieved August 4, 2003 from http://www.usoe.k12.ut.us/curr/ednet/training/resources/ednetRes/Role_Instructor.pdf
10. Lapadat, J.C.: Written interaction: A key component in online learning. Journal of Computer-Mediated Communication, vol. 7(4) (2002) Retrieved October 3, 2003 from http://www.ascusc.org/jcmc/vol7/issue4/lapadat.html
11. Mezirow, J.: Transformative dimensions of adult learning. Jossey-Bass, San Francisco (1991)
12. Moore, M.G., Kearsley, G.: Distance education: A systems view. Wadsworth Publishing, Belmont, CA (1996)
13. Philpott, M.: The Broadband Home (2006) [On-Line] http://www.becta.org.uk/research
14. Plater, W.M.: Chaos and community (1998) [On-Line] http://www.oaa.pdx.edu:80/oaadoc/OAANEWS/FallSymp/platerfull.html
15. Richardson, J., Ting, E.: Making the most of interaction: What instructors do that most affects students' perceptions of their learning. In: Paper presented at the 5th International Conference on Asynchronous Learning, College Park. MD (1999)
16. Sherry, L., Myers, K.M.: The dynamics of collaborative design. IEEE Transactions on Professional Communication 41(2), 123–139 (1998)
17. Slavin, R.E.: Cooperative Learning: Theory, Research and Practice. Prentice-Hall, Englewood Cliffs (1990)
18. Spiro, R.J., Feltovich, P.J., Coulson, R.L.: Two epistemic world-views: Prefigurative schemas and learning in complex domains. Applied Cognitive Psychology (Special Issue on Reasoning Processes) 10, S51–61 (1996)
19. Whittle, J., Morgan, M., Maltby, J.: Higher learning online: Using constructivist principles to design effective asynchronous discussion. Paper presented at the NAWEB 2000 virtual conference (2000) Retrieved April 8, 2002 from http://naweb.unb.ca/2k/papers/whittle.htm
20. Thomas, J.W., Funaro, G.M.: A multi-media, computer-based model for learner-directed, collaborative problem-solving. In: Woolf, B., et al. (eds.) Working Notes of 1990 Spring Symposium Series on Knowledge-Based Environments for Learning and Teaching, pp. 68–71. Stanford University (1990)
21. Verneil, M., Berge, Z.: Going online: Guidelines for faculty in higher education. Educational Technology Review 6(3), 13–18 (2000)
22. Walker, H.: Computer Supported Collaborative Learning. In: ITiCSE, Conference on Intergrating Technology into Computer Science Education. Upsala, Sweden (June1-5 1997)

The Role of Web-Based Learning Environments in Fostering Collaboration

Jasna Kuljis and Lorna Lines

School of Information Systems, Computing & Mathematics
Brunel University, West London
Uxbridge, UB8 3PH, United Kingdom
{Jasna.Kuljis,Lorna.Lines}@brunel.ac.uk

Abstract. Web-based learning environments are increasingly used to support lecture and seminar activities in blended learning courses and there is growing evidence to suggest a positive relationship between peer-to-peer learning activities and academic attainment. The role of web-based learning environments that are used as a learning-support tool is considered where the blended learning system goals of *social interaction* and *personal agency* within the context of web-based discussion boards are of particular interest. Using web-based learning community discussions as a case study we illustrate the role of personal agency and social interaction and how personalisation of the learning experience may be supported by the teachers. Our results are intended to inform teachers and the design of automated instructional agents, that will need to support personal agency and social interaction, to facilitate online student collaboration within the blended learning system context.

Keywords: Blended learning, collaborative learning, online discussion, discussion group development.

1 Introduction

The integration of web-based learning applications within traditional, face-to-face learning systems is growing rapidly with Higher Education Institutions in United Kingdom. Few guidelines are available within the literature to support such course designs and delivery due to the complexity of student and tutor needs, subject domain and the availability and maintenance of technological infrastructures. Rather, the pedagogical literature presents case studies that are considered to report 'best practice', usually providing an account of blended learning system designs that have been created on an ad hoc basis leaving the reader to discern the applicability of techniques and approaches in developing their own appropriate blended learning systems. Here, it is recognised that to provide a one-size-fits-all approach to the design of blended learning systems is inherently complex, and potentially futile, given the observed variety in blended learning sophistication. Our position is that although the one-size-fits-all approach is ambitious, as a learning community we may be able to guide designers of blended learning systems with specific learning components; a common trend in the current literature, whereby the focal point for discussion and

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 667–676, 2007.
© Springer-Verlag Berlin Heidelberg 2007

subsequent guidance may be, for example, the learning impacts offered by the use of web-based applications to support virtual classrooms [1] group work [2] and discussion boards [3].

Social interaction and personal agency are critical goals within any blended learning system. These goals may be supported through the use of web-based asynchronous discussion boards, whereby participation within the discussion is encouraged but not compulsory. Thus the goal of social interaction is promoted and supported and the goal of personal agency, through activities such as student choice and empowerment, is fostered. Of particular interests here, is how discussions are initiated, and personal agency developed.

To further the development of blended learning systems, this paper considers the role of asynchronous discussion boards within encouraged, but non-compulsory, student-to-student collaborations in problem solving activities. Having reported the benefits of adopting a blended learning approach to course design and delivery and identified the role of web-based discussion boards as an 'opt-in' personalisation feature, to support the blended learning goals presented by [4], [5] of *social interaction* and *personal agency*, are considered. Using discussion data collected from non-compulsory discussion board collaboration activities undertaken by a first year group of BSc (Hons) Computer Science and Information Systems students at Brunel University in the UK, the purpose of this paper is to identify discussion group development and a specific set of guidelines for tutors using discussion boards as a feature of their blended learning system approach.

2 A Blended Learning Approach

Blended learning systems integrate both face-to-face and online learning experiences and are inherently complex [6]. Student and teachers' needs, subject domain, and resources are all likely to impact the composition of blended learning systems and the achievement, or not, of enhancing the learning experience [6]. We take the position that the student should be the central component of any learning system, blended, or otherwise.

There is strong evidence to suggest that the success of blended learning systems lies in the integration of the online learning and face-to-face learning components during course design and delivery (see [5]). Good systems achieve the ultimate goal of enhancing student learning through utilising the strengths of face-to-face learning systems to counter weaknesses of online learning systems, and vice versa. For example, using a blended learning approach, face-to-face learning components are likely to include more group discussions and question and answer sessions between students and their teachers, which overcomes the relative weakness associated with isolation due to the lack of face-to-face communication activities. In contrast, self-test activities may be better suited to an online blended learning component as these could consume valuable time within seminars that could be more usefully employed to hold theoretical discussions or conduct practical activities.

Although 'integration' of the two learning system components (online and face-to-face) is critical, there are few guidelines within the pedagogical literature on how to do this. This may simply be due to the variety of courses, subject domains and disciplines. Due to blended learning novelty and complexity, current research is based

predominantly on case study reports for course designers to review and elicit salient points of best-practice for the learning needs of their students, the subject domain and their course management requirements.

However, although complexity in design and delivery is reported, the ultimate goals of blended learning appear to remain relatively simple and consistent. Osguthorpe and Graham [4] outline six goals of blended learning that permeate throughout the literature and are summarised as follows; *pedagogical richness* achieved by improvements to student learning; *access to knowledge* achieved through the accessibility of course content and materials that assist students' when demonstrating understanding, application and critical evaluation skills; *social interaction* which is measured as product of social operation through sharing and questioning learning experiences and activities with peers and teachers; *Personal agency* achieved by the student being supported in controlling and directing their learning experiences, leading the student to a sense of empowerment and heightened motivation; *Cost effectiveness* achieved by appropriate management of costs relating to a) teachers' time/effort and b) funding for the technological infrastructure and upkeep, and; *ease of revision* that can be measured against cost, time and human resources required to introduce and/or modify course content, materials and delivery.

The achievement of pedagogical richness within a blended learning system can be measured in improvements to student learning. Such improvements will likely include objective measures of student attainment and subjective measures that consider the students' personal learning experiences. The blended learning system goals of *social interaction* and *personal agency* are primarily based on student centred activities, whereas *cost effectiveness* and *ease of revision*, appear to relate directly to course management activities.

The lack of social interaction is frequently cited as a complaint offered by students undertaking distance learning programmes (see [7]) whereby face-to-face interaction (if any) is limited and social isolation is experienced. Within a blended learning system that comprises both face-to-face and online components it would be reasonable to assume that any sense of social isolation is reduced in comparison to purely online, web-based, distance learning courses whereby students are geographically dispersed and face-to-face events are not supported within course design and delivery. However, what is also unclear in blended learning system development is whether social interaction, a critical goal and a positive contribution to the students' learning experience (see [4]), is a necessary activity within the online component of a blended learning system. Due to the positive effects on the student learning experience, we take the position that social interaction should be encouraged, but not necessarily enforced, within any learning system and its sub-components. Although student collaboration and discussion has been observed to be of merit for student satisfaction and in some cases, improved attainment (see [1]), this is not, in itself, justification for enforcing collaboration activities; especially when the goal of *personal agency* is considered.

The goal of personal agency is to provide students with the opportunity to choose learning system activities that best support their individual needs. By offering a sense of control over their learning system choices, the student then 'buys-in' and takes ownership of their participation. This commitment and identification of their own learning needs is likely to lead to greater reports of student satisfaction and student

attainment as identified by Hiltz [1]. In practice, it is very difficult to meet all student learning needs while balancing the management of associated activities especially with large groups of students. However, with the additional tools that online learning applications now offer, such as virtual chat, discussion boards and video-conferencing, there is some opportunity within courses to support the goals of *student interaction* and *personal agency* simultaneously.

2.1 Discussion Boards: Supporting Student-to-Student Collaboration

The impetus on teachers who adopt a blended learning system is to strike the right balance between traditional face-to-face learning components and their choice of online learning components. Although discussion tools are offered as a positive feature of online learning tools, little is known as to how these discussion boards should be utilised to facilitate the goals of social interaction and personal agency. The literature tends to concentrate on the use of web-based discussion boards within the context of traditional distance learning courses that do not support face-to-face learning activities. Within the distance learning context, rather than discretionary use of discussion board features, the contribution to the discussion of each student may be monitored by the teachers, and/or contribution to the discussion may be compulsory (for example, see [7], [8]).

As supportive blended learning system features to be used to facilitate interactions outside of the classroom, we take the position that discussion board tools may be better placed as 'optional' learning system components that the student can opt-in to, rather than opting-out of. For a student to opt-out of using online discussion board facilities may be of equal detriment, if not greater, to enforcing interactions between students. By opting out of using the discussion board, the student is making a clear statement that they do not wish to participate. By allowing students to opt-in, they are less likely to fear reprimands from the teachers or their peers.

Thus, we liken this encouraged, but not enforced, use of an online discussion board within blended learning systems to wider ranging online applications, such as e-commerce sites and e-government sites that allow users to choose content and delivery design features to enhance their online experience.

3 Case Study

A case study approach was adopted using discussion board postings from a first year programming course *Construction of Programs* offered at Brunel University in the UK. This course is compulsory and the main emphasis of the course is object oriented programming principles using Java. The course employs the blended learning approach that combines formal face-to-face lectures and practical work undertaken in lab sessions together with extensive support provided via the WebCT learning platform [9]. We combined the strength of BlueJ, an interactive development environment designed explicitly as an environment to teach introductory object-oriented programming (hhtp://www.bluej.org) and WebCT.

There are 2 hour lectures plus 2 hours of structured lab activities a week over 24 weeks. The course is supported by a teaching team of tutors and graduate teaching assistants. Lectures are given in one session to all students enrolled on the course.

Students are divided into smaller groups and each has a 2 hour lab session per week. To support lectures and labs, we use Barnes and Kölling [10] which focuses on the general object-oriented and programming concepts while integrating this with the use of the BlueJ environment. The course has a strong web resource base and all course material is available on WebCT including coursework online submissions, regular tests and quizzes, and an on-line discussion forum.

Throughout the course students are required to perform a lot of practical problems as well as to complete associated tests assessing their knowledge. Each week students are given practical exercises (made available on WebCT) that address the material presented during the lecture for that week. Typically, most of the students have been working on the given exercises during their lab sessions in the same week. Lab sessions are supported by a group of tutors and teaching assistants available to help with the problems. Students are asked to undertake the corresponding lab test upon the completion of lab exercises. Each lab test is made available for two weeks and students were given freedom to choose when to undertake the test. Most of these tests (80%) were assessed and contributed 25% of the overall grade for the course. Students had information on which tests would be used for the assessment and how the tests would be graded and therefore put considerable effort to perform well on those tests. Students also had to put all acquired knowledge into practice through a substantial programming project which was assessed and contributed 25% towards their overall grade. The exam which contributes 50% towards the overall grade is a combination of multiple choice questions and problem solving questions to which students are expected to provide their own answers.

The on-line discussion forum on WebCT is a standard facility offered by WebCT in the form of threaded messages. It was used throughout the duration of the course to post notices to students. They were informed about its existence but no active effort was made to enforce any type of students' participation in the discussion forum apart from discouraging the use of email queries related to the course directed to individual tutors. Students were advised to post such queries to the WebCT discussion forum so that all students could benefit from the question/response and reduce the number of identical queries. Even though the members of the teaching team regularly check the discussion board and reply to posted questions, if or when relevant, student participation was neither monitored nor measured for any performance purpose. The decision to use the WebCT discussion forum was entirely the students'.

3.1 Data Collection and Analysis

WebCT offers course administrators an archive facility, whereby message boards can be downloaded and their content stored. The discussion board messages for 2003-04 were gathered via this archive technique and used for the subsequent analysis. All of the messages posted on the discussion board were subjected to a thematic analysis [11] whereby common themes based on, for example, message topic, function, and activity were identified within each message and grouped to obtain an understanding of the message-posting behaviours of both students and teaching team members. In total 726 messages were posted on the discussion board of which 612 messages were posted by students and 114 messages by the members of teaching team. Unsurprisingly discussion activity was observed to be at its most prolific immediately prior to the coursework deadline and examinations periods.

Table 1. Messages by themes

Theme	All messages	Messages posted by students
Organizational	329	249
Problem solving	397	367
Total number of messages	726	616

As an output of the thematic analysis, two major themes were observed as a) organization, and b) problem solving. Within the *Organization* theme nine sub-themes were identified and seven *Problem Solving* sub-themes were observed. The number of messages in *Problem Solving* theme messages slightly outnumbers the messages in *Organization* theme (see Table 1). The number of problem solving messages was steady with an upwards trend through the period until March when the number of messages soared due to increased coursework activity that was due at the beginning of April. Similarly just before the examinations.

3.2 Supporting Social Interaction and Fostering Personal Agency

In a face-to-face learning setting there should be no problems in supporting social interaction. After all, requirements of physical presence on campus, at least for scheduled activities, such as lectures and lab sessions, inevitably would develop a sense of belonging to a group by its members. So, the question we should address is what prompts social interaction in an online environment? We have seen in our discussion so far that students voluntarily participated in the discussion forum and that they posted 616 messages in the period of 11 months. There was no pressure to use the WebCT discussion forum, the participation was not monitored and, most notably it did not affect assessment. So, why did the students participate? Was their classroom / lab based social interaction insufficient; what were the lacking aspects?

We tried to answer these questions through an analysis conducted on all recorded messages in the WebCT discussion forum for the course. First, we observed that less than half the class participated and that only 98 students from the class of 195 actively participated in the discussion. What motivated participation and what were the reasons for 50% abstentions we can only speculate? Whatever the reasons, these non-participating students may have reaped other people's efforts. The majority of active students (60 students) posted fewer than 5 messages, and only a very small minority (17 students) posted ten or more messages. In the cases where students posted ten or more messages we observed the following pattern; the majority of students initially sought information from the teachers, then from each other and then they started providing help. One can assume that this indicates that their confidence, and in turn their *personal agency* grew.

Table 2. Distribution of problem solving messages

Subject matter	All messages	Messages posted by students
Coursework related	205	189
Lab exercises	51	47
Lab tests related	69	66
Exam related	38	36

The analysis of message categories indicates that assessment is the main concern of the most messages with the request for code or solution to the coursework project problem being the most frequently posted message. Coursework related organizational messages and general course related messages follow. Surprisingly the next most frequent category of messages is provision of code or solution as the response to a request and these were supplied for coursework related problems. This suggests that there indeed was a sense of community created by the discussion forum, as evidenced by the solicited and unsolicited help posted on WebCT. Even though students helped each other on a variety of messages we consider problem solving messages a primary source for observing the development of personal agency and table 2 illustrates the distribution of problem solving messages, whereas table 3 identifies the 'learning' role of the message posted.

Table 3. Students' collaboration on problem solving

Type of involvement	Number of messages
Seeking help	145
Giving help (solicited)	123
Posting solutions (unsolicited)	11
Active collaboration on problem solving	268

Unsolicited messages were interesting; one can either assume that students posting them were either altruists or so proud of their achievements. However, our analysis is more concerned with the solicited help. Some students were very pragmatic members of the 'silent crowd' who only used the discussion forum when that was for their own benefit. Several of them in the process also shared some of their knowledge. From our analysis we can argue that the discussion forum tool offered on WebCT was instrumental in the formation of an online community. Students who were actively engaged in shaping this community obviously felt as its members as their behaviour demonstrated. Some members developed into specific roles through the engagement in the community. In particular one author was soon recognized as a benefactor by the members of the online community who often addressed their queries to him.

4 Discussion

Our case study illustrates the use of a threaded discussion forum offered by an online learning environment WebCT. We wanted to assess whether such discussion tools could not only support social interaction but also foster the development of personal agency; two critical blended learning goals. The course used in the case study has a strong problem-solving orientation and students are assessed on a variety of problem-solving activities throughout the teaching year. The course taught students how to approach problem solving using the Java programming language. We felt that the combination of the complexity of the subject and the difficulty many students experience in learning programming together, with continuous assessment based on problem solving activities encouraged students to collaborate.

As already mentioned, in each week of term (course run over two terms) students were given a set of exercises to complete during the scheduled lab sessions and a test to attempt. Most students attempted all tests regardless whether or not tests were assessed. We deliberately did not provide solutions to lab exercises, lab tests, and to the previous years' examination papers to further encourage students to rely on their skills, to collaborate on problem solving activities, and be less dependent on their teachers. Members of the teaching team regularly monitored the ongoing discussions on WebCT and intervened when necessary. However, we did not reply to queries immediately unless the issues that rose required the response from one of the teachers. Most of the queries were left unanswered for a day or two so that other students might feel inclined to answer instead, which in most cases they did. We interfered only if the answers provided by other students were inaccurate, or if nobody made an attempt to answer, or sometimes to provide a broader view. That approach worked well and there were no complaints from students. The involvement of students in problem solving activity was gradual but towards the end of the course the forum was buzzing with constructive discussions.

Many students also worked together in small groups formed during scheduled lab sessions. However, such partnerships were not always successful, as sometimes groups formed by the academically weak students needed help from outside. Similarly, academically strong students often formed their own groups and there was not always interaction between these two types of groups. Given the blended learning environment and availability of real social interaction within the community created through attending lectures and lab session, the question is what prompted half of students to use the online discussion board to the extent that the development of personal agency could be observed. Some of the reasons are as follows:

- Many students work and have no time for social interactions, or have no flexibility with their free time;
- It is convenient to use the discussion forum, any time and anywhere (Internet);
- The immediacy of one-way communication – posting a message.
- In face-to-face interaction you need to know whom to ask which is not always obvious. In a virtual community questions are addressed to everybody who has access and chances are somebody coming with a solution to a problem are much higher;
- There is relative anonymity and hence less embarrassment for participants.

The above analyses complements the points made by [12] regarding asynchronous discussions. The idea of opting in to participating in discussion groups is supported by the trend towards learner control [13], where learners' choose the mechanisms, or rather, learning system sub-components.

There is evidence that the community was quite strong. It persisted well after the course was finished – it only stopped when students were denied access because of the need to prepare for the next year. What can we say about the 'silent crowd', the half of the students? There may be many reasons, such as shyness, anxiety, lack of confidence/time, information overload and/or poor motivation [14]. One of the challenges is to find mechanisms for increasing the proportion of participating students while providing the opportunity within a constructive, non-threatening environment to develop personal agency.

5 Conclusions

This paper considers the role of asynchronous discussion boards within encouraged, but non-compulsory, student-to-student collaborations in problem solving activities. The analyses of messages posted on the discussion board on WebCT throughout the duration of a programming course indicated that students were indeed collaborating on problem solving activities. Assessment was the main concern of most messages.

Based on our experience and the evidence from the conducted analysis we give the following recommendations:

- The tutor role within the discussion should be perceived as facilitator, not provider. Resist 'spoon feeding' as it does not help develop personal agency nor independence of thought.
- Encourage, but do not enforce participation in online group discussions. Do not monitor student participation for assessment purposes as additional 'pressure' to perform may detract from the development of personal agency.
- Do not respond to student queries quickly to encourage community development.
- Step in only when there is a problem. For example, when there is no response to a student query; a repeated request or students are suggesting inappropriate ways of solving the problem.

Although we support the notion of non-compulsory participation within online discussion activities, caution in the use of a number of optional face-to-face and online learning system components, may negatively impact the students' learning experiences. Kay [13] warns that by providing the student with too much control over their learning activities, we may induce a sense of overload to the point of distraction, thus making any choices that the student has made regarding their learning preferences redundant.

The results reported here are intended to inform teachers attempting a blended learning approach as well as informing the design of automated instructional agents to facilitate online student collaboration within the blended learning system context. Our tutor recommendations support, in part, the activities to be conducted by the automatic peer as described in [15].

References

1. Hiltz, S.: Impacts of College-level Courses via Asynchronous Learning Networks: Some Preliminary Results. Journal of Asynchronous Learning Networks (1997) available: http://web.njit.edu/~hiltz/workingpapers/philly/philly.htm
2. Gonzalez, M.G., Burke, M.J., Santuzzi, A.M., Bradley, J.C.: The Impact of Group Process Variables on the Effectiveness of Distance Collaboration Groups. Computers in Human Behaviour 12, 629–648 (2003)
3. Schellens, T., Valcke, M.: Collaborative Learning in Asynchronous Discussion Groups: What About the Impact on Cognitive Processing? Computers in Human Behaviour 21, 957–975 (2005)
4. Osguthorpe, R.T., Graham, C.R.: Blended Learning Environments: Definitions and Directions. The Quarterly Review of Distance Education 4, 227–233 (2003)

5. Graham, C.R.: Blended Learning Systems: Definition, Current Trends, and Future Directions. In: Bonk, C.J., Graham, C.R. (eds.) Handbook of Blended Learning: Global Perspectives, Local Designs, Pfeiffer Publishing, San Francisco (2004)
6. Garrison, D.R., Kanuka, H.: Blended Learning: Uncovering its Transformative Potential in Higher Education. The. Internet and Higher Education 7, 95–105 (2004)
7. Clark, J.: Stimulating Collaboration and Discussion in Online Learning Environments. Internet and Higher Education 4, 119–124 (2001)
8. Corich, S., Kinshuk, L.T., Hunt, L.M.: Using Discussion Forums to Support Collaboration. In: 3rd Pan Commonwealth Forum on Open Learning, Dunedin, NZ (2004)
9. Kuljis, J.: Orienting the Teaching of an Introductory Object-oriented Programming to Meet the Learning Objective. Journal of Computing and Information Technology. 12, 135–142 (2004)
10. Barnes, D.J., Kölling, M.: Objects First with Java. A Practical Introduction Using BlueJ. Pearson, Harlow (2003)
11. Owen, W.F.: Interpretative Themes in Relational Communication. Quarterly Journal of Speech 70, 274–287 (1984)
12. McNeil, S.G., Robin, B.R., Miller, R.M.: Facilitating Interaction, Communication and Collaboration in Online Courses. Computers & Geosciences 26, 108–699 (2000)
13. Kay, J.: Learner Control. User Modeling and User-Adapted Interaction. 11, 111–127 (2001)
14. Nonnecke, B., Preece, J.: Why Lurkers Lurk. In: Proceedings of Americas Conference on Information Systems. Boston (2001)
15. Goodman, B.A., Linton, F.N., Gaimari, R.D., Hitzeman, J.M., Ross, H.J., Zarrella, G.: Using Dialogue Features to Predict Trouble During Collaborative Learning. User Modeling and User-Adapted Interaction. 15, 85–134 (2005)

The Virtual Paddle: An Universal Interaction for Accessible Video Games

Guillaume Lepicard[1,2], Frédéric Vella[1], Nadine Vigouroux[1], Benoît Rigolleau[2], David Chautard[2], and Elian Pucheu[2]

[1] IRIT, UMR CNRS 5505, DIAMANT Team, Paul Sabatier University
118, Route de Narbonne, 31062 Toulouse Cedex 9, France
{lepicard,vella,vigourou}@irit.fr
[2] ENAC, 7 avenue Edouard-Belin BP 4005, 31055 Toulouse Cedex 4

Abstract. The video games are often inaccessible to disabled motor people of the upper limbs. The specialized games are limited choice of the set of accessible games. We present a new accessible technique of interaction: the Virtual Paddle. We designed it with a user-centered method. It is composed of three steps: questionnaire, brainstorming and paper prototypes. Next, we present the experiment to choose the optimized paddle. Finally we discuss the results and our perspective.

Keywords: Accessibility, software/virtual interaction techniques, video games, User-centered design, disabled motor people.

1 Introduction

In 2004, for the first time, the receipts of the video games exceeded these of cinema. 33 millions of console games and PC CD-ROMS have been sold. The receipts correspond to 1,454 billion euros against 1,134 billion euros for movies. (http://rss.zdnet.fr/actualites/informatique/0,39040745,39227871,00.htm?xtor=1).

However, many of these games require more and more efforts of learning for their handling. Indeed, the video games become increasingly complex. Most of the video games require a strong aptitude of reactivity to be controlled. Thus, popular video games are not accessible to disabled people. That contributes to those people's feelings of being excluded from the games performed by valid people.

Every disability has its limits for using a video game:

- Visually impaired people have difficulties to feel the colors and the shapes of objects of video game;
- Motor impaired people have limits to play through dynamic games;
- Hearing impaired people do not perceive the sounds of the game;
- Intellectual disabled and illiterate people cannot understand the rules and the goal of the games;

The disabled motor people present different behaviors during the use of a video game. The most frequent discomforts are: the muscular stiffness, the ataxia (muscular tone weak) and the deformation of the upper limbs.

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 677–686, 2007.
© Springer-Verlag Berlin Heidelberg 2007

We have defined four main groups of disabled motor people:

- Spinal cord injury: paralysis of the two or four members, generally due to a road accident, of sport, work or domestic. In certain cases, the people affected by these injuries can make use of some muscles of the arms and hands (amongst other things), which can allow small movements. In other more serious cases, the paralysis is total;
- Cerebral palsy: disorders motor leading to a disorganization of the movement and posture. These disorders are not related to the muscles themselves nor with the nerves which order them, but with a damage which has occurred in the central nervous system;
- Muscular dystrophy: cover various diseases neuro muscular of genetic origin. Duchenne muscular dystrophy for example, led to the atrophy of the majority of the muscles of the body;
- Locked-in syndrom: it is due to a vascular accident destroying the cerebral trunk. The consequence is a complete paralysis of the human body, except the eyelids. The eyelids make it possible to communicate with outside. Moreover the brain is intact.

The disabled people are often several disabilities. Some other problems such as spasticity (muscles stiffened permanently), dyskinesia (badly coordinated movements) and cognitive disorders penalize the player.

In this paper, firstly, we report the user-centered approach developed to design the Virtual Paddle for disabled motor people. Next, we will present a questionnaire concerning the accessibility of video games, and we will discuss the results obtained. We will present an experiment of the virtual paddle designed. Finally we discuss the preliminary results.

1.1 Game Devices

It is still difficult to find devices adapted to disabled motor people allowing playing video games.

The pioneer in this field was the NINTENDO society[1] which built in 1988 a device named “Hands Free”[2]. With this device the disabled motor people could play with the mouth. Some companies sold console games in which standard buttons of control can be replaced by others of bigger size. For example, Michael Andrews[3] built console games with switches[4]. This device was compatible with the Nintendo NES, SNES and later N64 games consoles. These devices are not enough adapted for the video games sold in trade. The reason is the speed of these games.

The games for Personal Computer are mainly played with the pointing device and keyboard. The reason is that these games require the combination of devices and the combination of several keys. These combinations can hardly be performed by people suffering from motor disorders of upper members.

[1] http://www.nintendo.com/home
[2] http://www.oneswitch.org.uk/2/hands-free.htm
[3] Pathways Development Group, Inc. (http://www.pathwaysdg.com/),
[4] http://www.oneswitch.org.uk/2/pioneers.htm

1.2 Video Games Adapted for Disabled People

There are video games entirely conceived by taking account of the disabled motor people. These video games are regulated by several parameters, such as the speed of the play, the number of lives of games heroes, the speed of action, the size of the interactive elements. It does ensure to the player to be more likely to enjoy pleasure while playing to video games.

The ICS–FORTH team has proposed the 'Universal Access" method [0] allowing to all disabled motor people to access to adapted video games. The game designers developed several technical interactions in order to adapt different modalities for each disability. UA Chess [0] and Space Invaders [0] are the games designed with this "Universal Access" method.

UA-Chess [0] is a chess game that can be played through a standard Web browser. This game can be used by different disabled people (e.g., low-vision, blind and hand-motor impaired). UA-Chess input (mouse, keyboard, microphone and switch(es)) and output (screen, Braille display, screen reader and loudspeaker) interface is adapted to player's profile.

Access Invaders [0] can be played in Windows and GNU/Linux. This game is used by different disabled people. This game is adapted to user's profiles. The game uses as input keyboard, mouse, vision–based gesture recognition, input musical, joystick, and one, two or three switches (or even more). The game uses as output spatial (3D) sound and a screen reader for blind people.

The "One Switch" games are games which we can be played with one only button. The OneSwitch[5] site proposes 70 games using this principle (only one button). The "Jetstream trader" game[6] uses a similar principle in order to top-up or top-down an air-balloon, horizontal displacements being carried out using a rush of air.

The principle of having only one button can be used for almost all actions (displacement, shooting, selection, jump, and so on).

On the other hand, it is not easy to conceive a whole game using only one button, and the more there will be actions to make, the more it will be necessary to memorize sequences of clicks. But it is possible to add buttons functioning on this model, and thus to manage to create a really captivating game.

These games stay marginal, and the disabled gamers want, like everyone, to try to play other games. Even, if they are not especially conceived for them. For that, they can use adapted devices.

The Table 1 summarizes the two techniques used for Universal Access to video games. The first one allows several inputs for each disabled people but the number of

Table 1. Accessibility solutions for video games

	Universal Access		
	Several inputs		Single input "One switch"
Game Name	UA Chess	Access invaders	Jetstream trader
Inconvenience	A few game exists		Limit of actions games, boring games

[5] http://www.oneswitch.org.uk/2/switch-downloads.htm
[6] http://www.brillsoft.com/Index_files/Page552.htm

games is limited (2 games). The second method uses only one input. This solution is not adequate to these people because the action games are limited. Behaviour reported by disabled gamers is that it is enough to bore their stiff. These two methods of "Universal Access games" cannot be integrated in games sold. One acceptable reason is the missing knowledge of the needs and requirements by the game designers.

We reported method and device that allow some video games to become adapted to disabled motor. Our objective consists in designing assistive technologies for video games. We have designed one assistive technology for games with user-centered method.

2 A User-Centered Design

The user-centered approach consists in three steps: questionnaire, brainstorming and paper prototypes.

2.1 Questionnaire

First, a questionnaire was designed to measure the accessibility of video games. Our objective was to know the users' profile and their difficulties on using video games. We wanted to have knowledge on their physic skills (members' movement, shaking, and so on) as well as to know their needs and requirements.

The questionnaire was published on the web during twenty days. We obtained seventeen answers but only fifteen were usable. The subjects are dispatched up according to their spinal cord injury disabilities: three with cerebral palsy, four with muscular dystrophy and seven with neuropathy. One third of these people have movement troubles.

The Fig. 1 represents the distribution of the movement difficulties of subjects who answered, according to the upper limbs that are affected. Each percentage represents the mobility rate of the articulations (see Fig. 1). According to the motor sensation of subjects the mobility of fingers and hand is more difficult than for the arm. The neck mobility is good.

This section will report both the use of the input device and the type of the keyboard. Six subjects use a virtual keyboard and nine use a classical keyboard. For

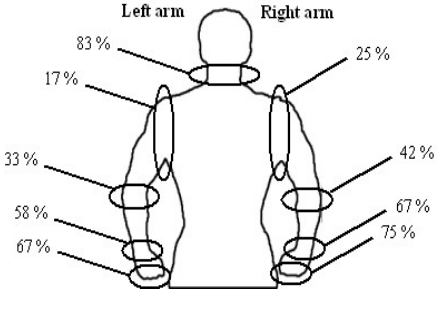


Fig. 1. Movements' difficulties

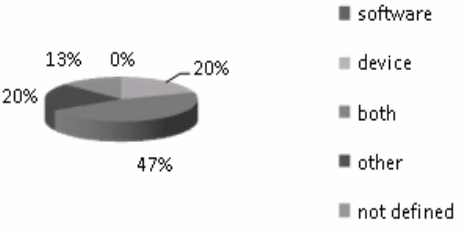


Fig. 2. Types of assistance tools

pointing device, nine people use a mouse, two a trackball and four another pointing devices (respectively a pad, a head-tracker and the keyboard arrows for two of them).

More interesting, a question concerns the wished assistance device (hardware, software or other) for video games (see Fig. 2).

The Fig. 2 shows that subjects (47%) would wish assistive technologies combining software components and physical devices.

2.2 Brainstorming

The questionnaire results lead us to do a brainstorming session. Two disabled motor (muscular dystrophy) subjects and five able-bodied subjects participated.

The brainstorming consisted in doing successfully the three following steps.

- A short presentation in three points: the main purpose of the accessibility project of video games, the questionnaire results and the devices available and potentially usable as switches, joysticks, and so on;
- The brainstorming itself where anybody gave his needs and requirements;
- A debriefing where the following recommendations were retained:
    - Supporting all video games;
    - Offering redundancy property [0] in terms of techniques of interaction; the subject must have the availability to realize an action in several ways, for example speech, click on a button, etc., see **Fig. 3** for a better understanding);
    - Offering the option of the key combination;
    - Designing a virtual paddle with only necessary buttons to use video games;
    - Allowing switches in different limbs.

Then, we realized architecture of the project components (see Fig. 3.).

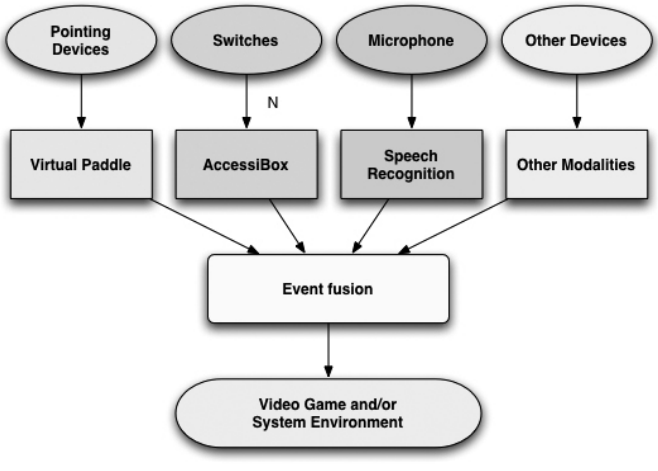


Fig. 3. Representation of the project components

There are three available modalities: gestural (pointing device), spoken (speech recognition command) and touch sensitive (set of switches). AccessiBox is a receiver of events from touch-sensitive contactors. Every interface (Virtual Paddle, AccessiBox, Speech recognition) transfers information to the "Event fusion" module. Currently, only one mono-modality is available: this means that a command can be expressed by one of these three modalities. This last module transfers the action to video games or/and to the System Environment.

In this paper, we focus the study on the virtual paddle. Firstly, we realized paper prototypes.

2.3 Paper Prototypes

Our aim is to maximize the movement time in order to realize actions with minimal latency and fatigue [0]. Each prototype of low reliability respects guidelines and principles recommended by previous studies and issues of the brainstorming:

- The distance of movement must be minimized (circle form is more efficient than lozenge form [0] (see Fig. 4 and Fig. 5));
- The most used keys are direction arrows: they will be set up in the center;
- The maximum number of action buttons must be seven plus or minus two [0].

Five paper prototypes were realized. Two prototypes are quite similar in layout and in the positioning of arrows (see Fig. 6 and Fig. 7). According to Fitts' law [0], the nearer arrow buttons optimize the movement time. Thus, we choose the prototype number five (see Fig. 8). The estimation time by Fitts' law calculated is 1, 056s (see Fig. 8) towards 1, 065s (see Fig. 7).

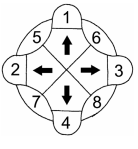


Fig. 4. Paddle 1

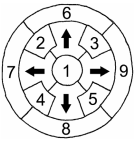


Fig. 5. Paddle 2

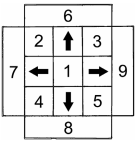


Fig. 6. Paddle 3a

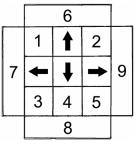


Fig. 7. Paddle 3b

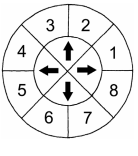


Fig. 8. Paddle 4

3 Experiment

The study was designed to access the paddle which optimizes the human interaction game.

3.1 Testing Environment

Four prototypes were built in Flash (see Fig. 9) after Fitts' law evaluation. The size of each paddle is 200x200 pixels in order to normalize distance. Paddle prototypes are built in Flash.

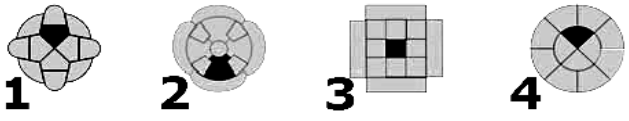


Fig. 9. Flash prototypes

User tests were realized on the web (http://jeuxethandicap.free.fr/).

3.2 Protocol

Each subject had to test each paddle with the same action sequences, for instance: up arrow, left arrow, right arrow, etc. The exercise consists in twenty selections by the usual pointing device of the subject (see Table 3). These targets become red when a click is required. The target move to another position after each validated click.

The subjects were asked to do the selection as fast as they could. They had to test the following sequence of paddles: N°1, N°2, N° 4 and finally N° 3.

3.3 Subjects

Nine users did the experiment: 3 with muscular dystrophy, 3 with spinal cord injury, one with cerebral palsy, one with neuropathy and another with genetic disease.

3.4 Data Collection

For each trial, the following parameters were recorded: the time in seconds between two clicks, the distance between two targets and the number of selection errors. All these data were saved in XML (eXtensible Markup Language) format. An analysis application was developed in order to play back user's actions and to understand the behaviour of the subjects (see Fig. 10, Fig. 11, Fig. 12).

Just after the test, a questionnaire designed as a debriefing, was elaborated to know from the subjects their preferred shape, the feelings of easiness and the fatigue induced by the couple (pointing device and paddle layout).

4 Results

Global results for the average time, number of errors and average distance are presented in Table 2.

Table 2. Results of the paddles

	Paddle 1	Paddle 2	Paddle 3	Paddle 4
Average time (s)	17. 09	17.162	16.195	15.729
Number of errors	2	1	0	2
Average distance (pixels)	1610	1660	1446	1594

Fig. 10 shows that the paddle 4 is more efficient than the others in regard to the time parameter. Moreover, standard deviation is shorter.

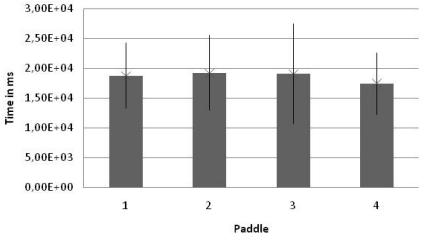


Fig. 10. Average time

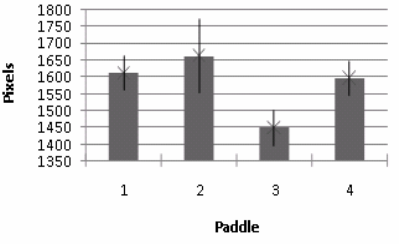


Fig. 11. Average distance

Fig. 11 presents the average distance gone all over the paddle according to the action asked. The paddle 3 is clearly more efficient than the others.

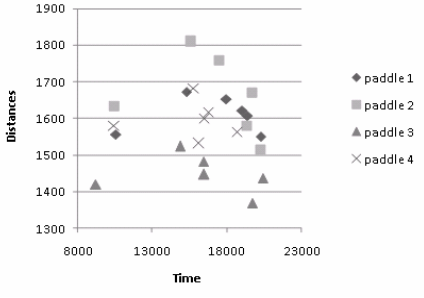


Fig. 12. Representation of paddles in the map: distance and time

Table 3. Devices of subjects

	Mouse	Track ball	Switch emulation	Head master	?
Muscular dystrophy	2	1	0	0	0
Spinal cord injury	0	1	0	1	0
Cerebral palsy	0	0	1	0	0
Neuro pathy	1	0	0	0	0
Other genetic disease	1	0	0	0	0
?	0	0	0	0	1

Fig. 12 confirms that paddles three and four are more efficient than paddles one and two. These results confirm the predictive time estimated with Fitts' law [2], as mentioned in 2.3.

Table 4 shows that the users prefer, firstly the paddle one and secondly the paddle four according to the two parameters: usability and layout.

A configuration tool was realized in order to connect the virtual paddle to "Event fusion" (see **Fig. 3**). This tool permits to define the commands linked to each button of the paddle. These commands have been designed to be equivalent to pressing a key on

Table 4. Notation of the usability and the paddle layout

	Paddle 1	Paddle 2	Paddle 3	Paddle 4
Usability	5	0	1	3
Paddle layout	4	0	1	4

a keyboard. The key pressing can correspond to: one action, a combination of several synchronic actions or a combination of several sequential actions. This functionality is vital for disabled motor.

5 Conclusion and Future Work

Thanks to user-centered design (questionnaire results and preliminary results of the experiment), we have chosen the paddle four (See Fig. 8) as input for a final design. This preliminary study confirms the importance of the layout for disabled people. In collaboration with a graphic designer, we designed the last available version for future experiments. We focused our thoughts on the shape, colours and icons semantic, news functionalities as transparency of the virtual paddle; etc.

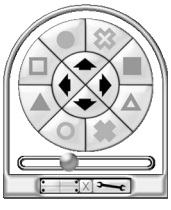


Fig. 13. Virtual paddle

We did not still evaluate this paddle (Fig. 13 in real situation). A new experiment is planned for measuring the efficiency of multimodality (possible choice of devices, see Fig. 3) for all types of video games, the correlation between the pathology and the input devices used. Our final aim is to compare performance and behaviour between disabled motor people and able-bodied people.

Acknowledgments. The authors gratefully acknowledge the valuable contribution of the subjects for the experiment. We also thank Jean-Léon Bouraoui for English feedback and checking.

References

1. Callahan, J., Hopkins, D., Weiser, M., Shneiderman, B.: An empirical comparison of pie vs. linear menus. Conference on Human Factors in Computing Systems, pp. 95–100 (1988)
2. Coutaz, J., Nigay, L., Salber, D., Blandford, A., May, J., Young, R.: Four Easy Pieces for Assessing the Usability of Multimodal Interaction: The CARE properties. In: Arnesen, S.A., Gilmore, D. (eds.) INTERACT'95 conference, pp. 115–120. Chapman&Hall Publ. Lillehammer, Norway (1995)

3. Fitts, P.M.: The information capacity of the human motor system in controlling the amplitude of the movement. Journal of experimental psychology 47, 381–391 (1954)
4. Grammenos, D., Savidis, A., Georgalis, Y., Stephanidis, C.: Access Invaders: Developing a Universally Accessible Action Game. In: ICCHP 2006 pp. 388–395 (2006)
5. Grammenos, D., Savidis, A., Stephanidis, C.: UA-Chess: A Universally Accessible Board Game. In: Salvendy, G. (ed.) Proceedings of the 3rd International Conference on Universal Access in Human-Computer Interaction. Las Vegas, Nevada, USA, Lawrence Erlbaum, Mahwah (2005)
6. Hinn, D.M, Bierre, K., Westin, T., Ludi, S., Ellis, B., Chetwynd, J.: Game Not Over:Accessibility Issues in Video Games. In: Paper presentation at the 11th International Conference on Human-Computer Interaction, July 22-27, Las Vegas, NV (2005)
7. Miller, G.A.: The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information. Psychological Review 63, 81–97 (1956)
8. Norman, A.D., Draper, W.S. (eds.): User-centred system design: New perspectives in Human Computer Interaction. Hillsdale (1986)
9. Stephanidis, C.: User Interfaces for All: New perspectives into Human-Computer Interaction. In: C. Stephanidis (ed.) User Interfaces for All - Concepts, Methods, and Tools, pp. 3–17 (2001)
10. Vella, F., Vigouroux, N.: Disposition spatiale des touches/caractères des claviers logiciels et fatigue motrice : premiers résultats expérimentaux. Handicap 2006, Paris, pp. 167–172 (juin 07-09, 2006)

Learning by e-Learning: Breaking Down Barriers and Creating Opportunities for the Visually-Impaired

Barbara Leporini[1] and Marina Buzzi[2]

[1] CNR-ISTI, via Moruzzi 1, 56124, Pisa, Italy
barbara.leporini@isti.cnr.it
[2] CNR-IIT, via Moruzzi 1, 56124 Pisa, Italy
marina.buzzi@iit.cnr.it

Abstract. Today e-Learning is an important educational tool with multiple benefits. However, to be suitable for a large audience the learning objects, courses, and other forms of e-Learning content should be easy to use for all students, regardless of disability. This paper describes problems often encountered by a blind person using screen reader and voice synthesizer when using e-Learning systems, and proposes guidelines for designers in order to develop more accessible systems.

Keywords: Accessibility, usability, e-Learning, blind.

1 Introduction

In recent years e-Learning has become a valuable tool for an increasing number of users. The benefits of this technology include (I) remote learning for students, (II) remote teaching for teachers living far from schools or universities, III) continuing education for adults.

A number of studies (e.g. [3], [11]) confirm that students appreciate the advantages of electronic materials in terms of portability and overall ubiquity, but they also value legibility, presentation and good design. There is also a clear demand for extra functions such as smart searches and dynamic indexing. Quality and the ability to provide extra facilities unavailable in paper textbooks are crucial for the future of electronic learning materials.

Distance learning can be a valuable opportunity for the visually-impaired and for other categories of disabled users, if suitable education methods and appropriate technologies are used. In particular, when a learning object or an e-Learning platform is accessible to but not usable by blind users, the benefits of distance learning methodology risk being worthless for them. Hence, when designing electronic learning materials (i.e. Learning Objects) and delivery systems (e.g. Learning Management Systems) the main issue concerns the users' abilities: it is crucial to identify the needs and requirements of the target community to create a system that fulfills their expectations. Thus a well-defined user profile is an essential component of the design process for the successful development of e-Learning Systems as well as Learning Objects.

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 687–696, 2007.
© Springer-Verlag Berlin Heidelberg 2007

In this preliminary phase of our research we analyzed user requirements and provide an overview of key aspects for the visually-impaired, highlighting unresolved issues and identifying future perspectives. It does not pretend to be an exhaustive analysis, but this initial investigation aims to provide an overview of the subject concerning people with a visual disability.

The paper is organized as follows: first we give an overview of studies in this field. Next we describe the main problems encountered by visually-impaired individuals when interacting by screen reader. We then discuss basic accessibility and usability issues related to e-Learning methodology and systems.

2 Related Works

In the last few years e-Learning technology has become an important research topic. Design and usability of e-Learning systems and objects is a primary focus of research in this field. However, interactive learning is still very difficult for blind persons who use assistive technologies such as screen readers and voice synthesizers. Various studies focus on the usability of e-Learning systems and some also include a general discussion on accessibility, but to our knowledge only a few specifically involve a study concerning the blind.

E-learning systems pose new challenges with respect to User Centered Design, where the target is a set of homogeneous users. Learner Centered Design must answer to the needs of multiple learner categories due to differences in learning strategies, know-how, experience in the learning domain, motivation to learn and user abilities.

In 1994 Nielsen proposed an informal method for evaluating systems and design usability based on a set of principles of usable software design (heuristics) [8]. A group of experts looks for violations of these heuristics, and thus the problems identified (a high percentage of the total) are fixed before the product arrives at the user. This method is cost-effective and easy to carry out compared to usability testing or other evaluation techniques. However, these principles were general, and in order to be applied to e-Learning applications and systems needed to be adapted to the specific domain. Squires and Preece in 1999 filtered usability heuristics throughout the socio-constructive theory and specified criteria ad hoc for e-Learning [10]. However various researchers argue the lack of accurate studies in this field. In [2] the authors take the first steps in defining a methodology for the rigorous evaluation of e-Learning applications, but accessibility for disabled students is not analyzed. Furthermore, Zaharias critically examined the usability of e-Learning applications and proposed a new usability measure: the student's intrinsic motivation to learn [12]. Developing a usability evaluation method based on a questionnaire he carried out two large empirical studies that showed the reliability of this approach. However, since it is a work in progress, details are not available regarding the accessibility section of the questionnaire.

Sloan et al. highlight limits and failures of the Web Content Accessibility Guidelines and propose treating accessibility with a holistic approach [9]. They believe that the goal of universal accessibility on the Web is inappropriate and that instead it is necessary to explore multiple routes to equivalent experience. As Kelly et al. argued, rather than demanding that an individual learning resource be universally

accessible, it is the learning outcome that needs to be accessible [7]. Based on user profiles, metadata and dynamic connection to resources, the experience of the user may be customized to match his/her abilities. We personally agree with this point of view and we believe that appropriate design is crucial for improving the accessibility and usability of e-Learning Systems. In [1] the study offers an example of designing UIs for the blind, applied to search engines.

Last, De Marsico et al. [5] define methodological guidelines of a joint design framework based on involving users with disabilities, as well as pedagogical experts, in the development process, believing that input of different know-how may enrich the quality of e-Learning applications, and provide a more satisfying learning experience. They also include two examples of building and providing learning objects accessible respectively to visually- (first example) and hearing- impaired students (second example).

Although the paper focuses on the needs of blind users, all disabilities should be considered when designing e-Learning applications. The criteria proposed ad-hoc for the blind must be integrated with others design guidelines. For example delivery of audio content should provide text and graphical content to supplement the audio channel (taking care to keep the content synchronized throughout the various media) and captions should be provided for any video incorporated into e-Learning applications.

3 Interacting Via Screen Reader and Voice Synthesizer

Designers of e-Learning systems must consider three crucial factors: usability, accessibility, and educational effectiveness. Consequently, the challenge is to design systems that are accessible to everyone and simple to use while maintaining pedagogical and educational efficacy. In particular, blind students may fruitfully utilize e-Learning systems if educational materials are accessible and learning paths can be tuned to the "rhythm" of the individual student.

When designing for blind users, it is necessary to consider the three main interacting subsystems of the Human Processor Model: the perceptual, motor and cognitive systems [4]. Sightless persons perceive page content aurally and do not interact with the mouse or other pointing devices since they only navigate via keyboard. This makes the "reading process" time-consuming and sometimes difficult and frustrating, if the contents are not designed with special attention to their needs.

The cognition part of the interaction is important, since many learning techniques are only relevant to people with good vision and may not apply to someone with a visual impairment. Thus alternative ways to deliver the same content should be provided. Furthermore, a blind person may develop a different mental model of both the interaction and the learning processes thus it is crucial to provide an easy overview of the system and contents.

Specifically non-visual perception can lead to many issues:

1. Content serialization. The screen reader reads the contents sequentially, as they appear in the HTML code. This process is time-consuming and annoying when part of the interface (e.g. the menu and/or the navigation bar) is repeated in many pages of the UI. As a consequence, blind users often have to stop the screen

reading at the beginning, and they prefer to navigate by Tab Keys, from link to link, or explore the content row by row, via arrow keys.
2. Content and structure mixing. The screen reader announces the most important interface elements such as links, images, and window objects as they appear in the code. For the blind user, these elements are important for figuring out the page structure. However, the actual reading process can overload the user, requiring considerable cognitive effort. If the table's content is organized by columns the screen reader (which reads by rows) announces the content of the page out-of-order, and consequently the information is confusing or misleading for the user. For the same reason, the table used for the layout must be avoided.
3. Lack of context. When navigating by screen reader the user can access only small portions of text and may lose the overall context of the page; thus it may be necessary to reiterate the reading process.
4. Lack of interface overview. Blind persons do not perceive the overall structure of the interface, so they can navigate for a long time without finding the most relevant contents. To resolve this problem it is necessary to structure the HTML code appropriately, add hidden information, and apply other features to improve navigation.
5. Difficulty understanding UI elements. Links, content, and button labels should be context-independent and self-explanatory.
6. Difficulty working with form control elements. The screen reader can handle the form elements by activating the modality – "form mode on" (e.g. by pressing the Enter key when the virtual focus is over a form control). In this situation, the screen reader informs the user by reading the sentence "form mode on". If the user forgets to switch on this modality, no typed text is written in a text field, no choice is selected from a combo-box, and some unwanted effects may occur since in modality "form mode off" any key pressure performs a different command (e.g. "t" moves to the following table and "h" to the following heading). Consequently, the user needs to re-start the operation.
7. Last, obviously, a blind person is unable to access multimedia content such as video streaming, video conferencing, and captioning.

4 Accessibility and Usability Issues in e-Learning

In order to create a distance learning course for all, the e-Learning system itself as well as the content should be adequate to users' abilities. E-Learning (literally electronic learning) is an all-encompassing term generally used to refer to computer-enhanced learning. It may include the use of web-based teaching materials and hypermedia in general, multimedia CD-ROMs or web sites, discussion boards, collaborative software, e-mail, blogs, wikis, text chat, computer aided assessment, educational animation, simulations, games, learning management software, electronic voting systems and more, as well as a combination of different methods. A particular contribution to the learning process in such an interactive environment is provided by

collaborative and cooperative technologies. For instance, the increasing trend is to create a Virtual Learning Environment (VLE) where students can collaborate and cooperate. Various studies to investigate the real effectiveness and efficiency of such environments have been conducted [6].

Pedagogically speaking, we should recall that the learning process is the main focus of an e-Learning-based education methodology. Students should be able to concentrate on contents and educational activities without being distracted by the system. That means that if all components are not accessible and suited to the user's capabilities, difficulties interacting with the system could distract the student from learning.

An e-Learning environment is basically composed of two main parts: (I) the integrated system – i.e. the container, which provides the contents and interactive activities to the participants; (II) the Learning Object (LO) - the educational contents provided through the e-Learning interactive systems.

Consequently, when designing an e-Learning environment suitable for various users including those with disabilities, two main aspects should be kept in mind:

- Accessibility and usability of the system, i.e. the "container" which provides contents and activities for the participants should be easy to use for all;
- Accessibility and usability of contents, that is the texts, documents and interactive tools used for educational purposes should be created according to various users' abilities.

If either of the two aspects is not accessible and usable, interactive distance education methodology risks being ineffective for people who are obliged to interact with assistive technology. This is particularly significant for blind users who interact by screen reader, which interprets interface content sequentially. Thus, it is important to consider both aspects when designing and developing a distance course.

4.1 Learning Management Systems

A Learning Management System (or LMS) is a software package that enables the management and delivery of on-line content to learners. Most LMSs are web-based to facilitate "anytime, any place, any pace" access to learning content and administration. Typically an LMS permits learner registration, delivery of learning activities, and learner assessment in an on-line environment. Both commercial and free platforms are developed. Open source and Web-based LMS software solutions are growing rapidly in the education and business worlds.

In addition to managing the administrative functions of on-line learning, Learning Management Systems also provide tools for delivering instructor-led synchronous and asynchronous on-line training (Learning Content Management Systems). An LCMS provides tools for authoring content as well as virtual spaces for learner interaction (such as discussion forums and live chat rooms).

Often, when considering accessibility issues only disabled students are considered, thus only LMSs are taken into account. Instead, people with disabilities may also be teachers, who must be able to create and manage the materials to use for distance courses. Thus LCMS authoring functionalities should also be accessible and usable by assistive technologies. In the following we generalize the discussion concerning

accessibility and usability features referring to both kinds of platforms (LMS and LCMS).

Since the user interface of an e-Learning system is usually based on Web technologies, at first glance it may appear sufficient to apply web accessibility guidelines. Actually, regarding the most widespread e-Learning platforms, some considerations can be made regarding technical aspects:

- Student tracking: provides information about the individual use patterns of the students in the course. Scorm-based platforms use javascript technologies for tracking the activities. For accessibility principles, the platform should offer the same functions when javascripts are disabled as well. Since most e-Learning systems – such as moodle (http://moodle.org/), atutor (http://www.atutor.ca/), and so on – are scorm-based, javascript is used. Alternative implementations should be investigated.
- Hidden frames: several e-Learning platforms use hidden frames for storing data to communicate to the Web server. Although for usability principles frames should not be used, appropriate design suggests using few hidden frames and positioning them after the frames containing the learning objects. In order to further simplify interaction, it is useful to give appropriate names to all the frames.

4.2 Learning Objects

A Learning Object (LO) may be defined as any entity, digital or non-digital, that can be used for learning, education or training (IEEE/LTSC). A Learning Object can be as small as a paragraph or as large as a complete online course and come in the form of HTML/Text files, simulations, JAVA, Flash, QuickTime movies etc. According to the format and complexity of a LO, various guidelines and principles should be applied to make them accessible and usable as well.

Apart from technical accessibility, we think that a specific issue should be considered when developing a Learning Object. Since the main goal of e-Learning methodology is education, a LO should be designed to exploit the user's abilities. For instance, for a sighted child an image or animation is probably more effective than a descriptive LO whereas for a blind child an audio object is much more effective than a textual one. Thus, the way a Learning Object is designed and structured is very important. In this case it could be made more effective by using different LOs for different users' abilities and skills. Further study is necessary in order to define a more appropriate approach for developing accessible and usable LOs.

Moreover, we think that most difficulties arise from simulation activities. Simulation systems and LOs can have various technical accessibility problems. A target study in this direction could prove useful.

5 Design Criteria

To make e-Learning systems accessible and usable, both the interactive system (Learning Management System) and the contents (Learning Objects) must be adequately developed. To this end, appropriate guidelines, requirements and suggestions can help developers keep the main issues in mind.

When designing an e-Learning system, the analyst needs to consider that in order to move quickly around the page content, the blind person prefers to use the Tab key to jump from one interactive element to another, rather than explore content via arrow keys.

Furthermore, objective difficulties may be reduced or eliminated by presenting the same content in aural form and providing multiple ways to navigate faster.

Last, when removing technical barriers for accessing contents, it is important to simplify user interaction in order to make the learning process more effective and satisfying.

5.1 Interface Overview

First of all, the blind person does not perceive the overall structure of the interface. Therefore it is important to structure the HTML code, defining logical sections of the interface.

Specifically, it is possible to group and structure (by heading levels) sets of homogenous text and elements in order to give the user immediately the idea of the interface, and enable him to jump rapidly from one section to another. For example, heading tags (<h1>..<h6>), if appropriately applied, are particularly useful since they are captured by screen reader and listed as an index, helping the user to "navigate" the interface.

Furthermore the use of some hidden information could help users to better comprehend the page structure.

5.2 Order of Content Blocks

Important content should be placed at the beginning of the interface code (i.e. the xhtml or html file) in order to facilitate sequential reading. Indeed, the position of the most important content is crucial for a blind user who usually wishes to explore the least amount of information necessary. Navigation of the container may become an obstacle to learning, so it is very important to place the most important element of the interface at the top of the content read by screen reader in order to access relevant content quickly. This feature was the most appreciated by totally blind users in our previous work [1]. The graphical interface instead must conform to the usual HCI criteria, arranging the content in a more appropriate way for visual elaboration.

5.3 Quick Comprehension of the Most Important Events

Aural feedback is very useful for the blind user since it associates a specific sound with a given situation (e.g. different tones may be associated with the success or failure of an operation) and simplifies interaction with form control elements.

The designer may include this feature in the e-Learning system interfaces, adding tones in some cases, such as:

- When the interactive form elements (edit field, radio button, or checkbox) receives the focus;
- When a lesson module is finished or a new module is started;
- To immediately inform the user of the success or failure of an operation (for instance the result of a self-assessment).

5.4 Simple Keyboard Interaction

Interaction with the system via keyboard is simplified if shortcuts and navigation via Tab keys are provided in order to jump directly to the most important elements of the UIs.

- Access keys must be associated with the most important interface elements offering a simple shortcut to the desired point (e.g. skip to the beginning of module, go to course index, move on evaluation assessment, and so on).
- Tab Keys define an order of visit of UI elements (e.g. links, objects). In this way, the order of "importance" associated with each element of the interface (i.e. the value of the tab access attribute) "drives" the user's navigation by Tab key. A lower tab-index value indicates greater importance.

5.5 Multimedia

Inability to access multimedia content such as video streaming, video conferencing, and captioning may be partially resolved by providing text equivalents, readable by screen reader. For example, supplementary content in an audio presentation can include the exact text presented as well as the description of relevant graphical content.

6 Discussion

E-Learning methodology can be an excellent educational opportunity for users with disabilities. A distance course offers students advantages such as flexibility and convenience, facilitation of communication between learners, greater adaptability to a learner's needs, more variety in learning experience with the use of multimedia and the non-verbal presentation of teaching material. Video instruction provides visual and audio learning that can be paused and reversed for watching again. Ability to access materials and the virtual environment both synchronously and asynchronously is a good opportunity for people who have difficulties in terms of time or space. This is particularly true for visually-impaired individuals who may have problems with classroom learning.

E-Learning in general education has been criticized since face-to-face human interaction with a teacher has been removed from the process; thus, some argue, the process is no longer "educational" in the highest philosophical sense. A blended methodology may be much more effective and efficient in terms of educational purposes, as observed by one of the authors of this paper in courses for blind students. Interacting with an e-Learning system is more complicated for a blind user than for a user with no visual impairment since interaction requires an assistive technology, which adds a level of complexity; thus an initial face-to-face phase may be necessary. Apart from this, e-Learning methodology represents a great opportunity for visually-disabled people, provided the interactive system is properly designed and developed.

The feeling of isolation experienced by distance learning students is often cited as a negative aspect, although discussion forums and other computer-based communication

can help overcome this and often encourage students to meet face-to-face and form self-help groups.

In the personal experience of this paper's blind author, we observed that blind users greatly appreciate mailing lists and, to a lesser degree, on-line forums. This is probably due to the fact that mailing lists have a sequential access; instead the forum interaction via Web interface may present usability problems.

7 Conclusions

In this preliminary work we have discussed accessibility issues for e-Learning systems (LMSs and LOs) and have proposed empirical principles for designers developing e-Learning applications in order to simplify interaction for a blind student or teacher.

We argued that both LMSs and LOs should be appropriately designed. E-Learning environments should be friendly and easy to use in order to reduce any possible negative effects on the learning process. However, rendering LMSs and LOs suitable for the abilities and skills of all users presents many challenges.

When defining the GUI it is fundamental to consider the needs of sighted users but the needs of blind users should also be taken into account when writing the UI code; furthermore the same information should be provided through both visual and auditory channels.

In conclusion, further research in this direction is needed. Future studies will investigate the empirical adaptability of an e-Learning system as well as suitable Learning Objects. Moreover, additional features of an LMS such as adaptability and customization will be addressed.

References

1. Andronico, P., Buzzi, M., Castillo, C., Leporini, B.: Improving Search Engine Interfaces for Blind Users: a Case Study. Springer UAIS Journal 5(1), 23–40 (2006)
2. Ardito, C., De Marsico, M., Lanzilotti, R., Levialdi, S., Roselli, T., Rossano, V., Tersigni, M.: Usability of e-Learning tools. In: Proceedings of the working conference on Advanced visual interfaces (2004)
3. Ardito, C., Costabile, M., De Marsico, M., Lanzilotti, R., Levialdi, S., Roselli, T., Rossano, V.: An Approach to Usability Evaluation of e-Learning Applications. Universal Access In. the Information Society 4(3), 270–283 (2005)
4. Card, S.K., Moran, A., Newell, T.P.: The Psychology of Human-Computer Interaction. Lawrence Erlbaum Associates Inc, New Jersey (1983)
5. De Marsico, M., Kimani, S., Mirabella, V., Norman, K.L., Catarci, T.: A proposal toward the development of accessible e-Learning content by human involvement. Springer, UAIS Journal 5(2), 150–169 (2006)
6. Eijl, P., Pilot, A., Voogd, P.: Effects of Collaborative and Individual Learning in a Blended Learning Environment. Education and Information Technologies 10(1-2), 51–65 (2005)

7. Kelly, B., Phipps, L., Howell, C.: Implementing a holistic approach to e-Learning accessibility. Retrieved at http://www.ukoln.ac.uk/web-focus/papers/alt-c-2005/accessibility-elearning-paper.doc
8. Nielsen, J.: Usability inspection methods. In: Heuristic evaluation, pp. 25–62. John Wiley & Sons, Inc, New York, USA (1994)
9. Sloan, D., Heath, A., Hamilton, F., Kelly, B., Petrie, H., Phipp, L.: Contextual web accessibility - maximizing the benefit of accessibility guidelines. In: Proceedings of the 2006 international cross-disciplinary workshop on Web accessibility (2006)
10. Squires, D., Preece, J.: Predicting quality in educational software: Evaluating for learning, usability and the synergy between them. Interacting with Computers 11(5), 467–483 (1999)
11. Wilson, R., Landoni, M., Gibb, F.: A user-centered approach to e-book design. The Electronic Library, vol. 20(4) (2002)
12. Zaharias, P.: A usability evaluation method for e-learning: focus on motivation to learn. In: Proceedings of CHI ’06 extended abstracts on Human factors in computing systems (2006)

Designing Games to Address 'Mute English' Among Children in China

Jason Li[1], Neema Moraveji[2], Jiarong Ding[3], Patrick O'Kelley[4], and Suze Woolf[4]

[1] Education Department, Brown University, Providence, RI, USA
[2] Center for Interaction Design, Microsoft Research Asia, Beijing, China
[3] School of Information, University of Michigan, Ann Arbor, MI, USA
[4] Microsoft Corporation, Redmond, WA, USA
jason_li@brown.edu, neemam@microsoft.com, dingj@umich.edu,
{pokelley,suzew}@microsoft.com

Abstract. This paper describes findings and design implications from user research intended to support software to address the Mute English phenomenon. This term refers to the many students in China whose English speaking ability lags far behind that of their reading and listening. Software designs for games are presented as a scalable solution to the problem. Games are a unique way to elicit speech input from children in a personalized educational setting.

Keywords: Design, ethnography, speech recognition, game design, China, Mute English, language acquisition.

1 Introduction

In an increasingly global economy, language-learning is a highly desirable skill. On one hand, face-to-face instruction with native speakers is ideal but expensive and difficult to scale. On the other hand, group classes sacrifice personalized learning. Technology, and software in particular, can alleviate the gap between the demand and the limited supply of fluent English teachers in a way that is cost-effective, personalized, and scalable. In rapidly developing China, the national language is vastly different from English and second-language acquisition of English especially difficult. Chinese college entrance exams taken by approximately 10MM students a year still concentrate on reading and listening, and one could safely argue that Chinese schools were 'teaching to the test.' But it is worth noting the new "speaking" and "writing" components to the TOEFL [ETS] required for graduate study abroad. The new TOEFL iBT—given for the first time in Chinese cities during fall 2006—devotes 25% of the test grade to speaking and 25% to writing skills.

"Mute English" is the term teachers and parents often use to describe the English that children learn in China: ability to read and write but an inability to speak. In the language acquisition literature this is known as the contrast between 'proficiency' versus 'communicative' measures. It is estimated that 60% of the current 230 million elementary school students in China are learning English [CENet], most likely Mute English, making the present problem an important one.

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 697–706, 2007.
© Springer-Verlag Berlin Heidelberg 2007

Our research project aims to understand this growing market, uncover user needs, and develop guidelines within which to design. We present the results here in the hopes of aiding other researchers to develop solutions to the Mute English problem; in particular, we consider the use of game design techniques to engage children in compelling, authentic spoken English scenarios.

2 Methodology

To better understand these education and cultural practices, we observed urban and rural Chinese classrooms, technology malls, and playgrounds. Additionally we interviewed teachers, parents, and children, conducted home visits, and tested English-language learning aides sold in China. These fieldwork exercises were punctuated with design sessions where we identified the user interface design implications of the findings.

2.1 User and Product Research

We conducted observations at 4 primary schools: in Beijing, 星星 Xing Xing School (private), 验 校 Zhengzhou Experimental School (public), 验 校 校 Zhengzhou Experimental School branch (private); in Lankao, one primary school (public), 1 middle school, and 考 Lankao Third High School. At each school we observed at least 2 English classes. We also conducted hour-long observations at 3 parks in Beijing: 武 (Xuan Xu Park), and the parks within 华 (Hua Qing Jia Yuan) and 来 (Xiang Lai Jia Yuan).

We visited 5 Chinese middle-class families, each of which has at least one primary-school-age child. Overall, we interviewed 11 students (Grade 1- Grade 6, ages 6-12), 6 primary school English teachers, 1 foreign English teacher and 10 parents.

To research similar products, an English teacher gave us a tour of the English education products sold at a large chain bookstore. We bought and tested the products she recommended as well as others that we found relevant to our designs.

2.2 Design

The first part of the design process connected research findings to design ideas on a mind map. During these brainstorming sessions, key research findings became major nodes on the mind map, and design ideas related to and inspired by these findings would branch off of these nodes. One design idea often inspired another, or was linked to another research finding, and we noted these on our maps.

After these brainstorming sessions, we placed these design ideas into context by embedding them within a storyboard. At first, storyboards were short and contained one design idea, but as we iterated the research and design process, storyboards began to grow in complexity and encompass a variety of designs.

3 Mute English

Both parents and teachers we interviewed spoke of the Mute English problem, sometimes translated as "deaf and dumb" English [Klein]. They acknowledged that English classes during school were doing a satisfactory job teaching reading, writing and listening, but not speaking. While children were able to read sentences from their workbooks, parents believed that the children's conversational ability needed work. Many middle-class parents sent their children to English tutoring sessions after school several times a week as well as for a half-day on Saturdays. If they could afford to, they went to one-on-one tutoring sessions. Despite all of this extra work, parents were still frustrated by the halting, heavily-accented English spoken by their children. During our research visits, the parents wanted their children to practice with our American interviewers.

Mute English is cyclic: people who've learnt Mute English inevitably pass it on to the next generation. Two parents we interviewed preferred not to practice English with their children lest they affect them negatively. Teachers devoted spare time to improving their spoken English, spoke humbly of their own English ability, and were open to receiving additional assistance. Software can present a scalable solution to this problem.

Aside from non-fluent teachers, there are two major causes of Mute English: the lack of opportunity for children to practice and their reluctance to do so. Software can provide safe and non-threatening settings for them to practice and motivate them by embedding the learning within a game. The secondary goal of perfecting pronunciation is not addressed because readily-available speech recognition engines lack the fidelity for pronunciation differentiation, particularly when children's speech and accents are involved.

The design challenge for addressing Mute English is to come up with creative ways to elicit spoken input. The goal is to have the child practice their conversational English in a free and casual manner. Furthermore, designs must fit local culture and practices and be compelling to children.

4 Findings and Design Implications

Rather than separating the results into user research findings and design implications, they have been grouped topically to emphasize the relationship between user research and design ideation. This coherence is a reflection of our process, in which each major fieldwork excursion was followed by a design session, leading to iteration that builds upon itself and upon new research. The observations and implications are relevant to a range of possible games.

4.1 Drills-Based, Competitive Pedagogy

Teachers in Chinese elementary schools, both rural and urban, use fast-paced competition to motivate students to learn English. As one teacher explained, competition is used to prevent the class from becoming a series of "boring" lectures. The method makes sense in a classroom because the words and phrases that students

know are basic and the class sizes are large, forty or more. A typical example: the teacher raises a large picture card while students in row one race to say the corresponding English word--"Art! Music! Math!" Fastest wins and the winners from each row gather at the front for a final round where the winner receives a sticker. Throughout the final round, the participants and audience are visibly excited.

Even though a tense competition may not be the best learning environment, it is a familiar one. Game designs can mimic this process by providing a private, single-player study mode followed by a social, multi-player test mode. In place of classroom, teacher and sticker, online competitions can be conducted through an intranet or the Internet but be scheduled by real-life organizations responsible for prizes.

4.2 Scaffolding Pronunciation

The term *scaffolding* here is used to denote supportive instructional assistance that can be gradually withdrawn as the learner makes progress [Vygotsky]. The scaffolding techniques we observed in classrooms were not novel [Gibbons], but they have been delineated below to present the Chinese context.

We observed:

1. Teacher alternatives to direct verbal translation: When teachers used unfamiliar English commands ("Please sit down") and their students did not respond, the teachers repeated with gestures as visual cues. Only in rare cases did the teachers resort to repeating in Chinese.
2. Written and spoken mixed languages: Teachers and textbooks used English and Chinese within the same sentences. When introducing new words, teachers often used Chinese for the 'glue' words of the sentence; with younger students, Chinese and English versions of the same word were repeated to reinforce the translation.
3. Piecing parts together: Class exercises often combined separate parts of a sentence. The teacher would point to an incomplete sentence—a cloze exercise--on the blackboard and say, "I went there by___?" The student would flip through a deck of answer cards, choose one, and apply it to the sentence, "I went there by *train*."
4. Designing software to mirror the scaffolding techniques of teachers makes the software feel familiar, as demonstrated by the numerous educational games that do so. Offering it at the precise level of the learner makes it personal and adaptive.

We created a user interaction framework using spoken and visual cues, English and Chinese subtitles with intermediate steps, and optional help. This system has a parallel in videogame design: adjustable challenge levels and a reward system.

4.3 Parental Involvement

Spurred by the hyper-competitive education system, Chinese parents worry about their children's education, reflected in the popularity of after-school tutors even for academically successful children. There is an additional layer of worriers: grandparents, less directly involved but with more free time. We saw neighborhood playground benches full of watchful grandparents and nannies. Each child of the one-child policy has possibly two parents and four grandparents to please.

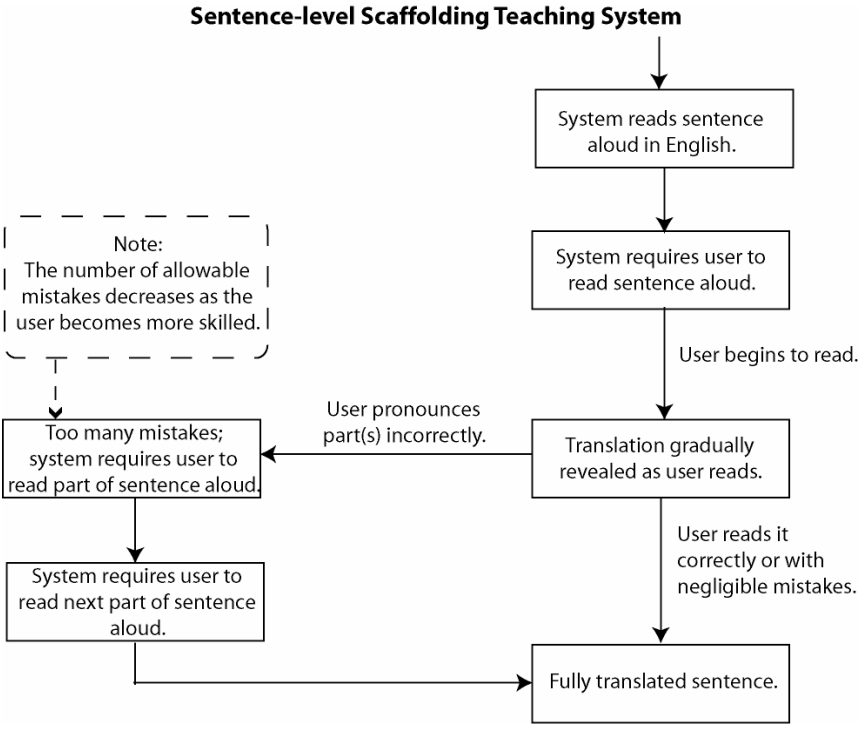


Fig. 1. Sentence-level scaffolding framework

Fig. 2. Parental supervision at a park within a housing complex

Conveying the progress a child is making to these six interested parties will be desirable. Software reinforces its pedagogical value when accompanied by learning progress reports. Parents and grandparents will be curious about both absolute and normative information, and there are a variety of methods by which this could be delivered to them, so long as the software collects the relevant data:

- from a third party online service that abstracts the child's progress and presents in a form understandable to the parent (though this requires the parent to have sufficient skills to know where and how to retrieve it)

- from the child's software activity itself, such as a "Reports" menu item; the child can also be prompted to print/e-mail an "award" page to show to a parent.
- from a "push" model, where a message and report is sent to a device or location the parent specifies as the one s/he wishes to receive progress information, which may be remote to where the child's activity is actually taking place
- or the software can directly address the co-located parent within the context of the child's activities.

These progress reports are particularly important because they provide assurance about the educational value of the game. Because there have been numerous reports in the Chinese media about videogame addiction, parents worry that their children will become obsessed with gaming and ignore their studies. Progress reports must then reflect a degree of academic rigor so that parents will encourage rather than forbid their children from playing such games.

4.4 Scheduling

The children we interviewed all had busy after-school schedules. Between tutoring sessions, art lessons and instrument practice, we found that their largest single block of free time typically fell on a weekday afternoon and lasted 2-3 hours. These few hours are currently spent on physical exercise, watching television and computer-related activities. Games must then be targeted towards the 2 to 3 hour block of free time, or be flexible enough to be played during breaks and unexpected free moments. Games might include many small, incremental steps (that can be saved session to session) that progress towards a larger game goal. Games should be approachable, with simple control schemes and integrated tutoring of the game rules so that children can immediately start playing and learn more subtle aspects of a game as it unfolds.

4.5 Foreign Avatar

A once-a-week class taught by a foreigner is the prevalent solution to the Mute English problem. Many parents wish their children to have more practice with fluent foreign teachers, even an unqualified foreigner. People believe that foreigners have a more standard and acceptable accent; one parent complained that people in other countries cannot understand speakers with Chinese accents. Teachers and parents both referred to an ideal "environment" for learning English, and foreigners are a key part of this ideal; the extremely expensive, sought after "American Village" summer camp is proof of this glamorized foreignness. Implicit is the idea that children must be immersed in an environment where they speak English all day, every day [Wighting].

Creating a software avatar is a way to bring the expensive foreigner to a wider audience than an exclusive summer camp, but with certain design implications. Children may be initially apprehensive about being observed by a foreign avatar, and it may be beneficial to begin a game with a Chinese guide. The foreign avatar can gradually take over until the Chinese guide retires. In this way, the child gradually comes to speak English to someone with no apparent understanding of Chinese.

4.6 Model Pronunciation

Children in China lack speech role models. Their parents were brought up with Mute English, and teachers prefer to use tapes in class instead of their own voices. Even English-speaking presenters on television, despite their fluency, have noticeable Chinese accents. The preferred accent on television is American, but in schools both British and American English are used in the tapes accompanying various textbooks.

Imagine a child avatar that is fluent in American English. He/she will act as a language partner and English-speaking model for the young user. The concept is potent because the child avatar can mimic pre-schoolers' and kindergarteners' tendency to enunciate by speaking slowly. Additionally, users may more readily identify with and be willing to mimic the mouth shape/actions of a child avatar in a pronunciation exercise.

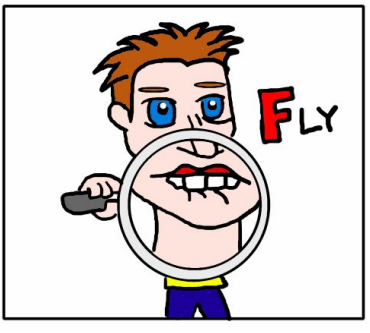


Fig. 3. Storyboard example: child avatar holds a magnifying glass up to his mouth to show the user how to enunciate the "f" in "fly"

5 Game Design Schemes to Elicit Speech

Because the goal is practicing oral English, children should first speak and then later tune the precision of their speech. Here are six design schemes that can be applied to a variety of games:

1. Shifting the cognitive load elsewhere: [Gibbons, p. 25]. Two examples are threading English words or phrases inside a song (standard practice in classrooms), or having children concurrently react to an on-screen task with their hands, where saying the English word or phrase is only part of the task.
2. Multiplayer competition: Like similar classroom activities, structure the task as a competition between peers, such as short rounds of "Word-Tetris," or a racing game where words act as fuel, or a cooperative construction game where pairs have to build a pyramid one word-block at a time, competing with other pairs of students.

3. A word a day casual game: There is a definite place in children's schedules to play short study-break games. Short frequent sessions of learning tasks, separated by time away, are associated with better retention rates than lengthy single sessions [Dempster].
4. Teaching the teacher: Chinese culture places a particular emphasis on respecting elders, and so children are often hesitant to speak freely in front of parents and teachers. To encourage speech, the foreign avatar should be more of a friend than a teacher. A role reversal in which children take care of the foreign avatar, and must learn English to do so, would be even better.
5. Immersive gaming: Immersion programs like "American Camp" don't allow students to use their native language.. Software games can immerse the child in an English-only environment, too, but with assistance and limited scaffolding given by icons, sounds, animations, etc.

6 Related Work

The use of games as a means of easing second language acquisition is an active field (e.g., [Cole, Johnson]). Speech recognition has been studied for such purposes as well [Hagen, Cole], but not to address the nuances of the 'Mute English' phenomenon, one where the emphasis is on encouraging conversation and correct pronunciation instead of learning new words.

The dynamics of game design as they apply to language acquisition have been studied in social- [Ang] and agent-interaction contexts [Brusk, Johnson, Kam, Wise]. Designing such games requires an understanding of non-command interfaces, as stipulated in [Nielsen].

The scaffolding literature is rich and the present work leverages only a foundational understanding of it, found in [Jenkins, Vygotsky]. The use of scaffolding has been applied in a classroom context but had not yet been formally transferred to virtual use.

Language-learning games that rely on speech recognition software are becoming feasible thanks to advances in speech recognition and synthesis technology. One notable example, Tactical Iraqi [Johnson], not readily available commercially, takes the form of an immersive (mixed with arcade) game for adults and is the result of numerous academic studies on game design and speech recognition. However, the teaching practices of the Chinese education system are different; compounded with Chinese cultural norms, it is difficult to transfer findings without a thorough understanding of the target users. Baldi [Cole] is a talking head conversational agent that uses scaffolding in the form of a "correct speech routine" that repeats until the child has spoken correctly. It uses a limited set of scaffolding measures to help the user.

6.1 Some Commercially Successful Technological Designs

On the whole, English-teaching technological aids are viewed with skepticism by teachers and parents in our pool of study participants. Despite this, there is much to learn from their designs. The most well-known line of aids, 读 (Readboy), is a tablet devices that reads out content from the national curriculum English textbooks

after the child enters the book and page number. One group of three teachers we interviewed believed only students "behind" in class should use this device; smart students would be able to absorb all necessary information during class time. Three pairs of parents felt similarly, and one offered that it was too " " (simplistic).

Fig. 4. Readboy's English language tablets

Sing-along video CDs and VCDs and interactive storybooks are more popular, as are online English-Chinese mixed newspapers, though these aids are all dwarfed in usage by in-person tutoring sessions and homework. These are all simpler than the Readboy, yet they seem to be appreciated for their material outside of the school curriculum. Ten of eleven children we interviewed said casual games are one of their favorite activities. During our home interviews, we observed their role as effective study breaks. However, none of the games they mentioned were tied to learning English.

There is a design conflict between fresh content favored by parents and children and time spent on the existing curricula. From our survey, the former is favored to the extent that innovative technologies are disliked if they only provide old content.

7 Conclusion

This paper presents motivation, data collection, and design implications for games that elicit speech input for Chinese children to learn English. Such games should first compel spoken interaction from their users, then fine-tune their pronunciation. Design implications have been presented to spur further research about the domain. Games that attempt to address Mute English could benefit specifically from a competitive atmosphere, an applied scaffolding framework, foreign child avatars, and parental involvement.

References

1. Ang, C.S., Zaphiris, P., Wilson, S.: Social Interaction in Game Communities and Second Language Learning. In: The 19th British HCI Group Annual Conference. Edinburgh, UK (2005)
2. Brusk, J.: Requirements for Spoken Conversational Agents in Games, Term Paper in the GSLT-1 course Speech Technology (2005)

3. CENet: China Economic Information Network, Beijing, China.(2005) http://www.cei.gov.cn
4. Chinnery, G.M.: Emerging Technologies: Going to the MALL: Mobile Assisted Language Learning. Language Learning and Technology 10(1), 9–16 (2006)
5. Cole, R., Massaro, D.W., Villiers, J., Rundle, B., Shobaki, K., Wouters, J., Cohen, M., Beskow, J., Stone, J., Connors, P., Tarachow, A., Solcher, D.: New tools for interactive speech and language training: Using animated conversational agents in the classrooms of profoundly deaf children. In: Proceedings of ESCA/Socrates sponsored Method and Tool Innovations for Speech Science Education (MATISSE) workshop, University College London, London (1999)
6. Cole, R., Carmell, T., Connors, P., Macon, M., Wouters, J., de Villiers, J., Tarachow, A., Massaro, D., Cohen, M., Beskow, J., Yang, J., Meier, U., Waibel, A., Stone, P., Fortier, G., Davis, A., Soland, C.: Intelligent animated agents for interactive language training. In: STiLL: ESCA Workshop on Speech Technology in Language Learning, Stockholm, Sweden (1998)
7. Dempster, F.N.: The Spacing Effect: A Case Study in the Failure to Apply the Results of Psychological Research. American Psychologist 43(8), 627–634 (1988)
8. Educational Testing Systems, TOEFL: Test of English as a Foreign Language. Princeton, NJ (2006)
9. Gibbons, P.: Scaffolding Language, Scaffolding Learning. Heinemann Press, New York (2002)
10. Hagen, A., Pellom, B., Cole, R.: Children's Speech Recognition with Application to Interactive Books and Tutors. In: IEEE Workshop on Automatic Speech Recognition and Understanding (2003)
11. Jenkins, H.: Motivation and Learning. Scholastic Press, New York (2005)
12. Lewis, J.W., Vilhjalmsson, H., Marsella, S.: Serious Games for Language Learning: How Much Game, How Much AI? In: Proc. of the 12th Int. Conf. on AI in Education (2005)
13. Kam, M.: MILLEE: Mobile and Immersive Learning for Literacy in Emerging Economies, Qualifying Exam Proposal
14. Nielsen, J.: Non-command user interfaces. Communications of the ACM 36(4), 83–99 (1993)
15. Klein, R., Rogers, P., Yong, Z.: ESL and Learning Technologies in Fast Developing China: Creativity and Innovation toward Beijing Olympics 2008. In: Proceedings of the 4th IEEE International Workshop on Technology for Education in Developing Countries (TEDC'06) (2006)
16. Vygotsky, L.S.: Thought and Language. MIT Press, Cambridge, MA (1962)
17. Wighting, M., Nisbet, D., Tindall, E.: Exploring a Summer English Language Camp Experience in China. A Descriptive Case. Study, Asian EFL Journal Quarterly 7(4), 85–108 (2005)
18. Wise, B., Cole, R., van Vuuren, S., Schwartz, S., Snyder, L., Ngampatipatpong, N., Tuantranont, J., Pellom, B.: Learning to Read with a Virtual Tutor: Foundations Literacy. In: Kinzer, C., Verhoeven, L. (eds.) Interactive Literacy Education. Lawrence Erlbaum Mahwah, NJ (In Press)

Identification of Threat of Isolation as a Result of Work with a Computer in Free Time

Teresa Musioł[1] and Katarzyna Ujma-Wąsowicz[2]

[1] Silesian University of Technology, Faculty of Organisation & Management
ul. Roosevelta 26-28, 41-800 Zabrze, Poland
tessa2005@wp.pl
[2] Silesian University of Technology, Faculty of Architecture
ul. Akademicka 7, 44-100 Gliwice, Poland
kwasowicz@poczta.fm

Abstract. The report aims to answer a question how we can reduce a threat of isolation created by using a computer in free time. In the first part of the report the reasons of creating a state of isolation during work with a computer in free time were presented. Then potential results of this threat in the area of optimalisation of decisions made by participants were discussed. In the application part of the report results of research concerning a state of awareness of isolation as a threat were presented. The research was carried out on a group of students, chosen at random. A sociometric tool in a form of a questionnaire was applied in the research. The construction of the questionnaire was based on weights of attribution of emotional state, which isolation is. In the summary the authors presented the ways and methods of reducing results of the threat, which is isolation in the process of working with a computer.

Keywords: isolation, computer, sociometric measure, reduction of threats, active recreation.

1 A Purpose of Undertaking a Problem of Isolation, That Is a Role of a Human in an Organization

The purpose of research, which result is this report, was an answer for a question if there is a problem of isolation caused by work with a computer and if a person subjected to 'the research' is aware of it. Furthermore, if such isolation exists, is its reduction in Polish society possible by active recreation and practicing sport in groups?

In the work process human activity contains rational and irrational elements. Irrational ones are purposes of the activity, whereas ways of achieving them are rational. Irrationality of these purposes lies in constant escalation of improvement a degree of consumerism, i.e. material needs. On the other hand non-material needs are given a role of a derivative, according to modern theories of management in market economy: i.e. first a profit and later the rest. An organization as a group of human individualities cannot constitute a monolith of rational actions. Irrational purposes of an organization are first of all time factors, it means the future determined by awareness of members of an organization [5].

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 707–715, 2007.
© Springer-Verlag Berlin Heidelberg 2007

Awareness doesn't exist as a separate function, having a structure and a location in a brain. It is not anything either internal or external. It is given so it meets the requirements of a phenomenon. Awareness is self-knowledge, which decides about an attitude of a human to himself. The more awareness, the more personality and will. A man who does not have will, doesn't have personality and self-knowledge, and in this way he proves to lack awareness. Lack of awareness is always a reason of evil and in human activity it creates a base for destructive processes in every organized sphere of life, first of all in an organization [6], [7].

In the industrial period an assembly line was a place of creation economic value of an organization. Today interpersonal communication with transmission of information about different values but first of all about global trade offer attracting consumer's attention is becoming such a place.

Every organization as a system, that is a group of elements and relations between them, is subjected to actions of powers of endo- and egzogenic characteristics. In order to realize its purposes, an organization needs to aim at balance between these powers, it means at a state of safety of an organization. The state of organization's safety is understood here as its ability to protect internal values from external threats, however it may be also the vice versa. It is related to a kind of these threats and their attribution in two aspects: objective and subjective. Risks existing in an organization are every event or a process undesirable from a point of view of undisturbed system operation. We can't talk about an isolated human or a situation. There is only a relation between a human and the environment – the relation which can be well defined by a word THREAT [9].

Isolation is a basic risk in the area of communication in process of every kind of work. Isolation, as it is defined by Norbert Sillamy [8] is 'a protective mechanism which aims at weakening perception by breaking it away from a context and its emotional basis'. It means that we deal here with emotional effort with a different degree of intensity, which depending on results of human actions, will have a sign 'plus' or 'minus'. According to Ruben Gallego [4], a bad system of organization, no matter what it concerns, dooms a human to isolation of any kind. The author states that isolation is misfortune and he even claims that every kind of isolation is a curse. It's difficult to disagree with this because a man is a social being and contrary to appearances emotional effort in communication with another person is essential, such as dynamic effort is indispensable for keeping correct figure. Isolation should not be confused with loneliness, particularly chosen one, because isolation is also a result of rational actions undertaken by natural, social and family environment on an isolated person.

2 Isolation of a Man Working with a Computer as a Threat

Threat created as a result of computerisation of every sphere of life – also free time, is first of all isolation of personal space. A man is assigned to NIPs, PINs, kodes, so he is identified by numbers. With help of his fingers he communicates with another person in the cybersphere, more and more controlled by other suitable and unsuitable people.

A man, who selectively receives information (limit of amount of information received per second), is not able to receive and process all information he would like to. Selecting requires decisions, i.e. a psychological act, and performing it requires a meaningful emotional effort. An examined man wants to be on his own and he thinks that he can't deal with this effort, and by some thought abbreviations he reduces not so much emotional, but rather intellectual effort. This is a simple way for limiting dialogue with another man, justifying pseudo individuality and negating group work. According to this, a man transmits more and receives less information. He starts to create his internal world, which characteristics obstructs reception of reality. It's simply better to dream than accept 'real' facts (are there any others?). It is a danger leading to all kinds of psychic disorders, because it influences lack of self-acceptance and in consequence to self-destruction in every form. All the more so because very often life in virtual reality deflates self-criticism and there follows subjective favourisation, known to industrial psychologists. Its variety is isolation in form presented by N. Sillamy [8].

3 Elements of Reducing Threats of Isolation, i.e. Active Recreation and Team Games

Sport is one of the most important spheres of human life. Practicing sports influences psycho-motoric development of a man and by positive rivalry it moulds his personality. It also allows for raising his self-esteem and self-realisation.

Awareness of values of active recreation (among others health benefits, not toxic rivalry, organization of time according to one's needs) and possibility of taking up sport (training disciplines that give pleasure) leads to reduction of civilization stress and aggression related to it.

Aware practicing sport is for many people an element of life that enables them to keep in good psychophysical condition. For instance, it makes it easier to lower emotional effort connected with studying and professional work. It works in case of low self-appraisal of one's appearance. It is recommended for health reasons for overweight people or in case of spine prophylaxis. Going in for sports outdoors is perfect for eyes and hygiene of sight [11].

In case of practicing sport motivative processes play an invaluable role. They can be defined as states of readiness to undertake a certain activity. We distinguish two kinds of motivation: internal (activation follows when a man aims at satisfying his needs, he is aware of a purpose that he aims at, and he does it voluntarily) and external (needs are created by applying punishments and awards). From a point of view of issues concerning active recreation, internal motivation is the driving force behind all activities. For some people it is a desire to have a beautiful figure (gym activities, aerobics classes), for others a desire to prove their courage (rock climbing on a climbing wall, evolutions in a skatepark), another group of people is attracted by a possibility to be a leader and to show off in a group (team disciplines) and some others by occasion to spend free time with friends.

In case of recreational sports essential issue is spontaneous self-organizing in more or less numerous groups, depending on a discipline. There is no need to appoint a leader, a trainer or a caretaker, although if there is such a person, all others often come

and practice because of his initiative. In team games, such as basketball, volleyball, football opposite teams should have one of two features: be 'balanced', if participants often change, or 'tight', it means that they practice regularly together and it results in understanding partners on a pitch or court. Good self-organization of such informal groups gives competitors a lot of satisfaction [10].

In Poland there is a problem of growing lack of interest in active recreation among people who complete primary education. One of crucial reasons of such attitude of society is deficiency in creating sports habits in a young man and suitable awareness for supporting them. Generally speaking, family, school and sports clubs should develop these elements.

In our report the following thesis was presented: if isolation caused by work with a computer in free time exists, it can and should be reduced by participation in active recreation and practicing sports in a group.

4 The Method of Measurement of Isolation Caused by Work with a Computer in Free Time

Three hundred twelve students of full-time and extension courses were subjected to research. The measurement of isolation was taken during semester classes (thanks to this students had direct contact with the researchers) in the period from 3rd to 27th January 2007. The students represented three faculties of the Silesian University of Technology: Faculty of Organization and Management, Faculty of Architecture and Faculty of Mechanical Engineering as well as Silesian College of Economics and Administration in Bytom.

The structure of examined sample is presented below in Table 1

Table 1. The structure of the sample accepted for measurement of isolation caused by work with a computer in free time (source: Authors)

Course mode	Age	Employment	Number of questionnaires
full-time	19-25	Working	110
		not working	6
extension	19-51	Working	29
		not working	137
Sum:			**312**

Table 2. The kind of the sample accepted for measurement of isolation caused by work with a computer in free time (source: Authors)

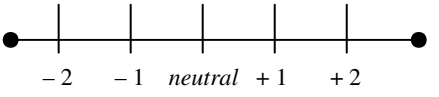
A	**Examined sample**	Full-time course not working	Aged 19-25	110 questionnaires
B	**Comparative sample**	Extension course working	Aged 19-33	110 questionnaires

Out of 312 received questionnaires, 30 items were filled in incorrectly. Out of correct 282 forms, researchers chose 220 questionnaires, according to criteria contained in Table 2

The measurement of a level of isolation was carried out with sociometric measure type 'Rating Scale Questionnaire' [3], constructed according to the following rules.

The measure contained:

- 10 open questions, a respondent was to answer them by choosing only one option,
- 5 possible answers for each of 10 questions, which were assessed in the scale from –2 to +2;



- the sixth answer – other comments – which was assigned to a proper scale;
- information gaps, such as age, course mode, date of measurement.

The interpretation of results of measurement of isolation was carried out by assessing every chosen answer for 10 questions in the scale from –2 to +2.

The questions included such areas as:

- ways of spending free time the most frequently and the most willingly,
- amount and quality of free time spent in front of a computer monitor,
- Awareness of risks related to work with a computer (among others isolation),
- Active recreation and participation in team games as a method of reducing risks related to work with a computer.

The presentation of results of isolation measurement lied in confirming or negating isolation during work with a computer in free time, coming out of a difference of positive and negative points in the examined sample. The answers described in the scale as 'neutral' (P_N) are considered as 'standard' behaviour and are not taken into account . Commenting on positive answers would be necessary if the difference occured to be 'negative'. The answers described in the questionnaire as positive defined the methods of reduction of isolation risk. First of all they include active recreation, participation in team games and social meetings in a group.

5 Analysis of Results of Measurement of a Level of Isolation During Work with a Computer in Free Time

Results of measurement of a level of isolation are presented in Table 3 and their graphic representation in Picture 1.

5.1 Interpretation of Results

Level of isolation $\Delta\ p_i$ was calculated as a difference between a number of points from answers to questions contained in positive scale and a number of points from

Table 3. Results of measurement of a level of isolation during work with a computer in free time (source: Authors)

Scale of Questions	*–2, –1, neutral, +1, +2*	
A Full time course, not working		
$\mathbf{P_{A-}}$ number of negative points	**– 2**	**– 1**
	number of answers 176 x (-2) = **– 352**	number of answers 289 x (-1) = **– 289**
$\mathbf{P_{NA}}$ number of neutral points	number of answers 266 = **neutral**	
$\mathbf{P_{A+}}$ number of positive points	**+ 1**	**+ 2**
	number of answers 205 x (+1) = **+ 205**	number of answers 168 x (+2) = **+ 336**
B Extension course, working		
$\mathbf{P_{B-}}$ number of negative points	**– 2**	**– 1**
	number of answers 161 x (-2) = **– 322**	number of answers 269 x (-1) = **– 269**
$\mathbf{P_{NB}}$ number of neutral points	number of answers 304 = **neutral**	
$\mathbf{P_{B+}}$ number of positive points	**+ 1**	**+ 2**
	number of answers 237 x (+1) = **+ 237**	number of answers 145 x (+2) = **+ 290**

answers to questions contained in negative scale in relation to: full-time course **A** (examined sample) and extension course **B** (comparative sample).

$$\mathbf{\Delta\, p_{i\,A}} = P_{A+} - P_{A-} = \{(+205) + (+336)\} - \{(-289) + (-352)\} = 541 - 641 = -\mathbf{100}$$

where:

$\Delta\, p_{i\,A}$ - level of isolation for full-time course, not working students,

$P_{A+} = P_{A(+1)} + P_{A(+2)}$ – summary number of positive points from answers.

$P_{A-} = P_{A(-1)} + P_{A(-2)}$ – summary number of negative points from answers,

$$\mathbf{\Delta\, p_{i\,B}} = P_{B+} - P_{B-} = \{(+237) + (+290)\} - \{(-269) + (-322)\} = 527 - 591 = -\mathbf{64}$$

where:

$\Delta\, p_{i\,B}$ - level of isolation for extension, working students,

$P_{B+} = P_{B(+1)} + P_{B(+2)}$ – summary number of positive points from answers.

$P_{B-} = P_{B(-1)} + P_{B(-2)}$ – summary number of negative points from answers,

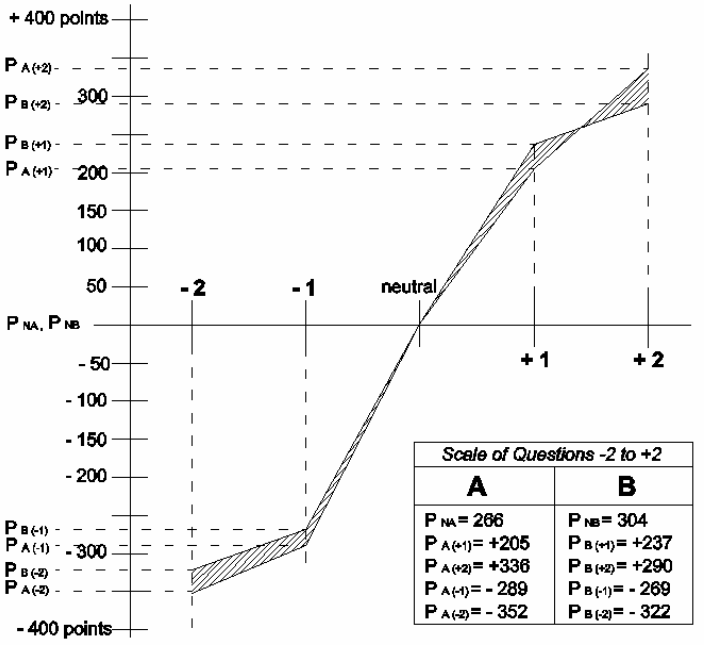


Fig. 1. Graphic representation of results of measurement of a level of isolation caused by work with a computer in free time (source: Authors)

Looking at the above data we come to a conclusion that the level of isolation of full-time students $\Delta p_{i\,A} = -100$ is higher than the level of isolation of extension students $\Delta p_{i\,B} = -64$.

While analyzing the reasons of isolation, number of points for answers in the negative scale and in the positive scale of full-time and extension students were taken into account [Table 3]. The conclusions are following:

1. The full-time students are not as aware of risk of isolation as the extension students ($P_{A-} > P_{B-}$);
2. The full-time students prefer active recreation and team games to a larger extent than the extension students ($P_{A\,(+2)} > P_{B\,(+2)}$);
3. The extension students are more aware of risks caused by work with a computer and they can reduce them better than full-time students ($P_{NB} > P_{NA}$). For this reason they are in favor of family and social contacts ($P_{B\,(+1)} > P_{A\,(+1)}$).

6 Final Remarks

- Growing meaning of knowledge and new economy of global society requires access to information and we can't withdraw from using more and more advanced information technologies, computer technics and software as well as hardware [2].

- The research of isolation risk caused by work with a computer in free time confirmed the thesis that computer users are not aware of this risk. However, instinctively they feel that active recreation and team games are the methods which help not only to relax but also to reduce the effect of isolation.
- In spite of rich sports base, a problem is a limited inclination of students for regular and also spontaneous practicing sport. One of essential reasons of such a situation is neglect in the period of bringing up a young man, because in this time follows a process of transformation exuberant willingness to spend free time in active way into a habit. The habit, which later appears to be very desirable.
- Widening educational market in Poland will require a transition from learning methods 'face-to-face' to methods 'e-learning', in which isolation will be not only a threat but also a social problem. While reducing intellectual effort and treating work with a computer as a reduction of emotional effort, such a situation may be a reason of degeneration of interpersonal communication, indispensable for team work.
- The research concerning a threat of isolation caused by work with a computer in free time, confirmed assumptions of the authors that students cannot manage health risk, connected with hygiene of studying and of working [1].

References

1. Błach, A., Musioł, T.: Issues of Safety and Higiene of Work in Learning Standards of Higher Education Institutions. Industrial Pedagogy Towards Problems of Work Protection. Publishing House of the Institute of Exploitation Technology in Radom pp. 155–160 (2004)
2. Charytonowicz, J., Sztajkowski, K.: Computer Environment and Automation in Building Designing Process. In: Stephanidis, C., Jacko, J. (eds.) Human – Computer – Interaction, vol. 1, Lawrence Erlbaum Associates, London (2005)
3. Charlton, S.G., O'Brien, T.G.: Handbook of Human Factor Testing and Evaluation. Lawrence Erlbaum Associates Publisher, Mahwah, New Jersey, London (2002)
4. Gallego, R.: I Am One of You. Tygodnik Powszechny weekly N°26' (2006)
5. Musioł, T.: Organization as a Phenomenon. Joint publication: Organization as Utopias. In: Rąb, J., Broda, J., Musiał, J.(eds.) The Society for Development of Polish Science, Zabrze (2004)
6. Musioł, T.: The State Ergonomics Consciousness Employed a Like Index the Quality Safety of Organisation. In: Stephanidis, C., Jacko, J. (eds.) Human – Computer – Interaction, pp. 1401–1405 Lawrence Erlbaum Associates, London (2003)
7. Musioł, T.: The State of Ergonomics Consciousness of Participants of Didactic Process in Individual Assessment. In: Stephanidis, C., Jacko, J. (eds.) Human – Computer – Interaction, vol. 1, Lawrence Erlbaum Associates, London (2005)
8. Sillamy, N.: The Psychological Dictionary. Libraine Larousse. Książnica, p. 107 (1994)
9. Sienkiewicz, P.: The safety of systems. Transformations - interdisciplinary magazine. The Educational Foundation Transformations and The L. Koźmiński Centre for Reasearch on Evaluation and Forecasting. Warsaw, pp. 303–312 (2005)
10. Ujma-Wąsowicz, K.: Outdoor Sport in the City. Development Trends. Public Space of Contenporary City. Cracow University of Technology, Kraków, vol. 2, pp. 303–312 (2005)

11. Ujma-Wąsowicz, K.: Ergonomics Factors of Quality of Active Recreation Environment. Conference materials from 34th all-Poland Ergonomic Conference OKE, together with 4th International Conference 'Man-Science-Environment' MSE 2006, organized by The Commision of Ergonomy PAN o/Wrocław, Wrocław University of Technology, National Labour Inspection, PTErg o/Wrocław, Medical Academy in Wrocław (in print 2006)

Reaching Beyond the Invisible Barriers: Serving a Community of Users with Multiple Needs

Dianna L. Newman and Gary Clure

University of Albany/SUNY, Albany, New York USA 12222
dnewman@uamail.albany.edu, gclure@uamail.albany.edu

Abstract. This paper discusses a four-phase model for evaluating multi-media learning materials that emphasizes the diversity of learners and variations in instructional needs and user characteristics. The authors begin with an overview of the model, supporting evidence for its use, and key characteristics of users supported by each of the phases. They then focus on results of a current use that emphasized stage four, real-time usability, and show how they were able to document that the models under review met the needs of diverse learners and varied instructional strategies.

Keywords: Human-computer interactions; user characteristics, effect of use; collaborative learning.

1 Introduction

The role of the student in technology-supported learning is no longer viewed as that of a passive recipient, but instead, is perceived to be the key that unlocks the process of development and use [1,2]. Earlier technology transfer practices based the development of computer assisted instruction and technology supported learning on the assumption that "technology is better" and assumed that we only needed a few research studies to find clear evidence of technology's impact [3]. Most of the studies that were conducted as part of this development were based on elementary information processing theories and compared alternate modes of delivery (e.g., computer vs. in-person) or alternative sequencing and/or degrees of use [4]. The emphasis of both development and research was on how to use technology to transfer or deliver new technology and the development of best practices in design and development. Results of these studies did not meet expectations of those who were designing and developing technology supported curriculum and instruction; in most cases, there was little, if any, clear evidence that instruction with technology resulted in greater cognitive outcomes than did other forms of in-class or distance education [5,6]. What did become clear, however, was that the individual learner played an important role in the process. Even when assuming an objectivist, teacher-centered approach to learning, it became clear that human- computer interactions and subsequent learning were in some way related to key learner characteristics such as learning style, gender, familiarity with content, technology anxiety, and motivation to learn. In sum, good design, that is "making technology use better," required attention to learner variations. At the same time that technology-supported curriculum was

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 716–727, 2007.
© Springer-Verlag Berlin Heidelberg 2007

expanding to include the dynamics of the user, the role of the learner in the overall process also was expanding. Constructivism and student-centered learning, coupled with social cognitive psychology, was identified as supporting the acquisition of many of the expected outcomes of technology-supported transfer of knowledge [7,8]. Using the assumptions of Vygotsky, Bruner, and other social psychologists [9], it was believed that not only did different learners bring their own characteristics to the process but that good technology-supported models would allow for instant scaffolding of materials that would assist in developing individual patterns of learning. As a result, designers and developers had to create technology-supported materials that would be adaptable to different styles of learners as well as different levels within a particular style. This broader definition of human computer interactions is now an important part of both design and use.

2 Purpose of the Study

The purpose of this paper is to present an evaluation model that is effective in documenting a constructivist, student-centered approach to technology-supported curriculum and instruction design and key user variables that should be considered as part of the process. Using examples of multi-media materials developed and piloted for pre and initial engineering classrooms, the paper discusses why we need to expand our boundaries of human computer interactions to include the needs of multiple users as well as some of the ways we can document the utility of the technology supported materials for varying types of users. A focal point of the presentation is findings related to stage four of the model, “real time usability” and key user variables that have been documented as being important at that stage.

3 The User Centered Model

If learner characteristics are important factors in the development of technology-supported instruction, it is necessary that these characteristics be reflected in the design and evaluation of modules that support learning. The challenge for IT managers is to lead development teams in designing technology-supported systems of learning that will facilitate the acquisition of factual and structural knowledge and its meaningful integration into the cognitive frameworks for learners from different backgrounds and with different degrees of prior knowledge [10]. As a result, the optimal goal is to structure the external conditions in a manner that will facilitate internal learning and information processing within each learner’s zone of proximal development or scaffolding level. To meet this need, Newman [11] developed a multi-stage process for team leaders to use in designing, validating, and documenting IT learning systems. Each stage of the model involves acknowledgment of users' characteristics, their prior levels of knowledge, skills and experiences, and the external, contextual variables that influence the formation of new learning systems. The model was used to develop and validate approximately 100 modules in multiple settings across multiple content domains for multiple types of users. Extensive notes on the process were developed for each stage and module. A meta-analysis of the

process across these users resulted in the identification of four phases involving assessment and inclusion of user characteristics [See 12]. The first phase identified by the meta-analysis is related to *content review*. Designing modules for learning systems typically requires use of experts in the field who are extremely knowledgeable about the content being created. Thus, the first phase recognizes cultural and contextual variables consistent with the biases of the designer and developer; it consists of an external validation of the module content. In the user-centered model, the design team utilizes experts representing theory and practice. Content area faculty and field based professionals review the instructional content and the proposed use of the content in various applied settings. The second reoccurring phase found in the meta-analysis consists of the need for an *instructional technology review* by experts in information systems design. The critical factor in these reviews is whether or not the technical design of the learning system acknowledges learner variables. Initial reviewers who are part of the design team work with project staff on a regular basis, providing formative feedback on user-centered needs. In addition, external reviewers who are specialists in instructional design and technology review and rate each system and/or module, assessing the relevance and flexibility of the presentation mode for different types of learners, and the flow of instruction and learning. *Usability evaluation* (phase three) specifically assesses the modules for relevance across different learning styles, prior experience with content, and previous experience with the specific mode of technology-supported system that is being developed. At least four types of learners should be involved in this process, those who are novice or experts in the content area and those who are novice or expert in the technology mode. In addition to being observed in the use of the module, participants also should be encouraged to "think aloud" as they go through the module, that is, to verbalize freely what they are thinking throughout their experience. The goal is to 'get inside the learners' head' and elicit what they were doing and why; this progression through the module along with the "think aloud" comments is videotaped and used to determine sequences and patterns of module use by different types of learners and the ability of the module to be effective for varied types of users. Where possible, variations in learning style also should be recorded. The fourth phase that consistently yields relevant information about users' characteristics involves direct observation of module use in *"real-time" settings*. This phase provides project staff with formative feedback on the utility of the modules in varied learning settings and work contexts. Procedures include observation of the use of the module, surveys of users, interviews with randomly selected learners, and follow-up interviews with instructional staff and developers.

4 The Program Under Study

4.1 The Modules

A multi-year project[1] involving the design, development, and evaluation of student initiated, student-centered learning tools utilizing innovative web based curricula in the pre-engineering and engineering domain serve as the background for this

[1] For more information on this project see http://www.academy.rpi.edu

investigation. The goal was to develop modules that would assist in presenting or creating knowledge in a manner that would allow for immediate transfer to practice in applied settings while motivating and rewarding learners as they progressed through the modules. Factors that needed to be considered as part of the development included students' gender, English as a primary language, prior use of technology and modules in learning settings, and variations in prior knowledge and experience. The modules also had to meet the needs of differing instructional strategies and teaching settings. To facilitate the development of the modules, a team consisting of discipline-based faculty, professionals from the specific content field, and specialists in learning, instructional design and evaluation, and technical developers was created. A formative, cyclical process of development, assessment, and evaluation was established that allowed the team to share information on an on-going basis, use discussion and feedback to make decisions and refinements in the process and the products, and share findings of intermediate outcomes. Provided in Table 1 is a partial summary of the modules under review and their ultimate uses in the classroom.

Table 1. Models of Module Utilization

Observed Classroom	Mode of module utilization	Activities during module utilization	Duration of module utilization
Amplifier Circuits			
Classroom 1	Embedded instruction	Guided inquiry	45 minutes
Classroom 2	**Summation**	Free exploration	25 minutes
Classroom 3	Summation	Demonstration	10 minutes
Classroom 4	Summation	Guided Inquiry	20 minutes
Power Factor Correction			
Classroom 1	Summation	Free Exploration	10 minutes
Classroom 2	**Embedded**	Demonstration	20 minutes
Classroom 3	**Summation**	Demonstration	30 minutes
AC Voltage Regulation and Efficiency			
Classroom 1	Embedded instruction	Guided inquiry	15 minutes
Filters CAD			
Classroom 1	Summation	Guided Inquiry	30 minutes
Nodal Analysis			
Classroom 2	Summation	Demonstration	10 minutes
Classroom 2	Summation	Free exploration	20 minutes
Classroom 3	Summation	Free Exploration	40 minutes
Pole Zero			
Classroom 1	Embedded instruction	***0Guided inquiry***	10 minutes
Classroom 2	Summation	Demonstration	10 minutes
Classroom 3	Summation	Guided inquiry	30 minutes
Classroom 4	Summation	Guided inquiry	25 minutes
2nd Order Systems			
Classroom 1	Summation	Demonstration	5-6 minutes
DC Circuits			
Classroom 1	**Advanced Organizer**	Free exploration	20 minutes
Exploring Connectivity K-12			
Classroom 1	Embedded instruction	Guided inquiry	30 minutes

4.2 Data Sources and Methodology

A series of *content reviews* was conducted for each module by team members who were experts in the material and by potential faculty users and advanced graduate students. The team conducted an initial review at the time of design (e.g., while "storyboarding") and later by independent faculty users and advanced students. The

final review was conducted during "real-time" use by students and faculty as part of the final evaluation effort. *Educational technology reviews* were conducted for 12 modules to assess their instructional principles by two experts in the area of learning cognition. The majority of the designed modules focused on the topics and materials covered at the beginning level of engineering undergraduate classes. *Controlled usability tests* (n=12) were conducted on three sample modules to determine their usability for students of differing content and computer ability levels. Each usability test consisted of participants working through a computerized module: one beginner computer user/beginner content, one beginner computer user/advanced content, one advanced computer user/beginner content, and one advanced computer user/advanced content. The purpose of the usability test was to demonstrate the level of computer ability and content knowledge required to successfully utilize the modules.

Multiple data sources and sub-studies were conducted to document *real-time usability*, the focal point of this paper. Paper-pencil surveys were distributed to and collected from undergraduate students who have been exposed to the computerized instructional modules (n=371). Survey respondents were enrolled in introductory Engineering undergraduate classes in 2005-2006. Surveys provided data concerning the level of utilization of the computerized instructional modules, and student perceptions of their affective and cognitive learning. Follow-up individual interviews were conducted with students (n=21) to determine perceptions of the availability, implementation, and utility of the computerized instructional modules. These interviews were conducted in-person at the conclusion of the class sessions. Observations (n=28) were conducted in classrooms to assess the setting and environment of module use, available technological support, and the participants' interaction with the instructional technology, instructors, and peers during module utilization. In addition to this naturalistic data collection, a series of causal comparative and quasi-experimental studies were conducted to test for relationships of module use and learner characteristics.

5 Findings

5.1 Phase 1 and 2: Validation of Content and Instructional Design Integrity

The modules exemplified many of the best practices in multi-media development that facilitate use by diverse learners with multiple needs. Content validation studies indicated the modules successfully presented information that could be used to provide supplemental or alternative content coverage as well as key concepts. Each module not only addressed the concept under study, but also provided examples of real world applications. A brief introduction summarized content and potential instructional applications including multiple ways in which instructors and learners could interact with the concepts. Both stand-alone use and integration-into-classroom instruction was facilitated by the sectioning of learning activities so that the material could be used in a continuous sequential manner or in a user/instructor selection approach. Modules also allowed for free exploration; students could take different paths with each that allowed them to explore some concepts more in depth than others. Many of the units also overtly supported constructivist learning by providing users with suggestions on how they could manage their own process through the

modules. As can be seen in Table 1, this flexibility allowed the modules to be used in studio classrooms, for demonstrations/lecture, as part of laboratory projects, and to support distance learning and study outside the classroom.

All modules involved the student in a number of reflective thinking activities, ranging from periodic question-and-answer exercises to direct opportunities to manipulate objects on the screen. Multi-media were incorporated in the forms of graphs, charts, diagrams, pop-up windows, and manipulative (applets). In addition, real-world examples were included. In some cases, graphical organizers and/or concept maps provided alternative representations of the structure of the module and related materials. Each module could easily be used to engage students in open-ended complex problem solving allowing them to develop and provide solutions with the opportunity to comment on the process. In addition, interactive modules were 'playful,' encouraging students to revisit content from different perspectives, and to take responsibility for their own learning. The modules also provided multiple opportunities for students to assess their own strengths and weaknesses.

In addition to interacting with the content, the modules provided students with multiple explicit opportunities to interact meaningfully with each other and/or the instructor. The inter relationships between the illustrations, content material, applets, and practice materials provided settings that supported collaborative work and discovery learning. The modules also provided exploration exercises that could result in collaborative decision-making and promoted peer teaching activities.

Basic tenets of development also were met: modules were carefully screened to eliminate culturally or gender biased language that could offend learners; the screens were simple, balanced, and easy to us; and text was visible, legible, and formatted for meaning. Finally, the differing kinds of content and activities were distinguished from each other in a consistent fashion with graphs, charts, and diagrams emphasizing important themes.

5.2 Phase 3: Controlled Usability

Controlled usability studies also documented that the modules met the needs of diverse users. Each type of user accessed information in a different pattern but was able to accomplish the expected tasks. Participants with high levels of technical skills and content knowledge could easily navigate within looking at both content and sequence; because they were familiar with the content, they attached meaning to the module structure and looked for their own pattern of information. Novice users, on the other hand, used the module in a very sequential pattern, going through each stage as presented and not moving forward until that phase was mastered. Content expert-novice module users spent more time looking at the layout of the module content and developing a cognitive structure for use; they then moved to those sections that provided them with the content they wanted to access and skipped over uses that were not of interest or were areas where they already understood what would happen. Module expert-content novices reviewed the layout of the module first, tried a few sample experiments to make sure they understood the command structure, and then began to sequentially go through the content. Participants, who were experts only at content but not module use, found that they could increase their knowledge base, but did initially experience difficulties navigating within some specific modules. Overall,

use indicated that if allowed to work through the module in a self-selected manner, all could complete the tasks and develop their own knowledge base.

5.3 Phase 4: Real Life Usability

As noted in Table 1, faculty utilized the modules in different ways: as an advanced organizer, embedded into particular class activities, or for the purposes of summation and review. In the advanced organizer mode of module utilization, the module was typically introduced at the beginning of the class to provide a framework for the session that followed. In the embedded instruction mode, the instructor would usually embed the module into the lecture to illustrate particular concepts or into activities as a problem-solving tool. In the summation mode of module utilization, the module was usually introduced at the conclusion of the class to provide another application of the material covered in class, or as a means to assist students in their homework assignments. In a typical session, the faculty would introduce the module as a supplementary material, which would reinforce the material covered in class. Instructions of how to navigate the module were thoroughly provided, and numerous applications of the modules were examined through examples. Students were expected to engage in a variety of activities in the process of module utilization; these included free exploration, guided exploration, or use of the module to complete an assigned activity. In many of the instances, students were given an opportunity to freely explore the module capabilities and draw conclusions of the modules specific applications. The modules also supported more traditional instructor–centered modes of teaching when use was restricted to instructor demonstration only. The module also supported variations in time of use; the duration of module exposure ranged from 10 to 45 minutes depending on the instructional messages that were conveyed. More complex modules incorporating multiple concepts required significant investment of class time and were used throughout consecutive sessions.

The majority of the students rated the different modules utilized in instruction favorably in terms of relevance, usefulness and organization of the information, and opportunities provided to practice course content; however, there were variations in the students' ratings based on key user characteristics. As indicated in Table 2, 95% of the students reported that the information presented in the module was relevant to the course content; 86% of students indicated that the modules provide opportunity to practice the course content, and 85% percent of them noted that the information presented was useful with regard to the real practice. Some variations by student characteristics were evident. Males tended to find the modules to be more useful in regard to practice than did females; autonomous users also noted them to be easier to understand and more related to course content. Collaborative learners, however, found the information to be more relevant; this latter response may be related to gender of subjects and should be investigated further. These findings were supported during interviews; several students indicated that most of the modules were easy to understand and self-explanatory. Students also pointed out that the modules served as a good reinforcement of content that had been taught and helped them to review the course material for exam purposes. Many indicated that the effective use of the modules was facilitated by the quality of the instructor-student interactions and the opportunities provided to students to interact with peers. These factors also increased their interest

and motivation. The observational data gathered confirmed that the use of the module increased student interest and tended to increase their motivation to stay on task. Additionally, observers noted that the use of the modules tended to add to the class dynamic and facilitated the effective use of group work. In interviews and on surveys, students indicated that the use of the modules facilitated collaborative group work; 67% of the students indicated that module utilization facilitated the process of collaboration with peers. Working with fellow students in solving problems was perceived as not only increasing the level of interactivity in the session but also as enhancing students' understanding and comfort with the material covered in class.

Table 2. Perceptions of Relevance By Student Characteristics Percent of Agreement

Type of Use	Gender		English Primary		Prior Use		Current Use		Total
	Male	Female	Yes	No	Yes	No	Auto	Co	
The information presented is relevant to the course content.	96	92	97	89	95	98	**93**	**98***	95
The information presented is useful with regard to course content.	92	88	91	96	91	94	**89**	**84***	91
The module provides opportunity to practice the content.	**88***	**75**	87	83	85	87	85	88	86
The information presented is easy to understand.	84	82	84	81	84	83	**88**	**81***	84
The information is well organized.	86	84	86	85	86	89	87	85	86
The information presented is useful with regard to real practice.	**86***	**76**	84	87	84	87	83	86	85

% agreement on 6 point scale (responses to three levels); * more than one standard error of measurement difference.

Data also indicated that diverse user needs were met, but differentially, when specific module and course goals were targeted. The modules were perceived as having helped the students to visualize problems, develop problem-solving skills, apply course content to new problems, and recall course content. Students indicated also that they would like to receive more explanations and guidelines on the appropriate use of the modules. As noted in Table 3, the majority of the student participants positively perceived the use of several modules with regard to specific learning outcomes. Specifically, 85% of the participant students perceived that the modules helped them to think about problems both in graphical and pictorial ways. On further query via interviews, a majority of the students commented on the graphic/visual component of the module, indicating that it was an aid to learning. Some of the students noted that the ability to see the layouts of the circuit, and the capability of changing the layout graphically, helped them to compare and understand concepts better. When queried as to the direct relevance to course content, 79% agreed that the modules helped them to better recall course content; during interviews, students reported that the modules reinforced the material covered in the classroom even if used before content presentation. Additionally, 72% of the students noted that use of the modules helped them to develop skills in solving problems with regard to the specific aspects of course content. Although both groups were positive about module use, those who used it in a collaborative setting had greater perceptions of its ability to help construct or reinforce specific content. Evaluator observations indicated

that students in these settings not only used the module but also verbalized the findings in order to work with a partner. This additional reinforcement and extension of learning may be a facilitator to long-term learning that is not available with autonomous learning.

Table 3. Perceptions of Effect on Specific Learning Outcomes by Student Characteristics Percent of Agreement

Type of Use	Gender		English Primary		Prior Use		Current Use		Total
	Male	Female	Yes	No	Yes	No	Auto	Co	
Helped me think about problems in graphical or pictorial ways	85	82	84	87	85	86	**79**	**89***	85
Recall course content	79	80	79	78	79	78	**75**	**84***	79
Develop my skills in solving problems in the course content	72	73	71	73	72	73	72	72	72

% agreement on 6 point scale (responses to three levels); * more than one standard error of measurement difference.

As noted in Table 4, effect of module use on general learning or course-wide goals are less emphatic. Between one-half and three fourths of students perceived positive outcomes in this domain but degrees of support were noted by student gender, English as their primary language and their use in autonomous or collaborative settings. Overall, 76% of the students perceived that they were better to apply course content to new problems, and more than two thirds of the students (68%) stated that the use of the module helped them to develop different approaches of solving problems and transfer knowledge and skills to problems outside of the course. During interviews, many students reported that the module helped them to understand course material; a few students indicated that the module had the potential of going beyond the course content. One of the features that students indicated they liked was the ability to input any number and observe the output immediately. This allowed them to try out different inputs thereby providing opportunities to practice content. Observations of classroom usage of the computerized instructional modules showed that the modules engaged students in problem solving tasks, encouraged students to revisit content from different perspectives, and helped them to take responsibility for their own learning.

Table 4. Perception of Effect on General Learning Outcomes by Student Characteristics Percent of Agreement

Type of Use	Gender		English Primary		Prior Use		Current Use		Total
	Male	Female	Yes	No	Yes	No	Auto	Co	
Apply course content to new problems	**78***	**67**	**74**	**84***	75	82	**73**	**80***	76
Develop different ways of solving problems	**66***	**77**	67	73	66	77	**63**	**71***	68
Transfer knowledge/ skills to problems outside of the course	67	63	**65**	**73***	66	69	**64**	**68***	66
Improves grades	58	57	59	55	59	51	**61***	**55**	58

% agreement on 6 point scale (responses to three levels); * more than one standard error of measurement difference.

The majority of students perceived module usage as increasing their knowledge base and understanding in the specific content area. Additionally, students were more active and engaged during module usage than during the lecture portion of the class, and students' confidence of using computerized instructional modules increased as a result of module usage. As noted in Table 5, students agreed that the modules met their learning needs (77%), assisted in increasing their knowledge (73%), enhanced their confidence in the content area (67%), and motivation (64%). Important variations were noted by student gender, English as primary language, prior use with modules and autonomous/collaborative uses. Classroom observations noted that when students used the modules working in teams they were highly engaged and motivated. During interviews, those students for whom English was not their primary language reported that they preferred the hands-on experience more than classroom teaching, as did students with no prior experience. During the interviews, some students voluntarily suggested that the modules be used as advanced organizers, noting that if they had been introduced in the class prior to going through all the content material, they would have grounded later learning of theoretical content. Amount of time was noted to be important; classroom observations showed that some students had a difficult time completing the module activities during the allotted time; many wanted more time to free-explore the concepts and develop their own pattern of interactions.

The modules also were reported to have a positive impact on the students' learning affect (See Table 6.) Overall, students reported that use of the modules positively affected their confidence in their knowledge of the content (81%); their interest and motivation (66% and 68%, respectively) and attitudes related to their ability to

Table 5. Perception of Cognitive Benefits by Student Characteristics Percent of Agreement

Type of Use	Gender		English Primary		Prior Use		Current Use		Total
	Male	Female	Yes	No	Yes	No	Auto	Co	
The module suits my learning needs.	79*	63	76	80	78	76	78	78	77
My knowledge has increases as a result of the module	**74***	**65**	**71**	**83***	**71**	**81***	**70**	**76***	73
My confidence in the content area has increased because of the module	67	67	**66**	**85***	**65**	**76***	66	67	67
The module content is motivating.	64	66	**63**	**74***	63	68	**62**	**67***	64

% agreement on 6 point scale (responses to three levels); * more than one standard error of measurement difference.

Table 6. Perception of Effect on Learning Affect by Student Characteristics Percent of Agreement

Type of Use	Gender		English Primary		Prior Use		Current Use		Total
	Male	Female	Yes	No	Yes	No	Auto	Co	
Developed confidence in content area	**81***	**73**	78	84	79	84	**79**	**83***	80
Developed interest in content area	66	67	63	82*	66	64	**63**	**68***	66
Developed attitudes of self-direction and self-responsibility	62	61	**60**	**71***	62	59	**59**	**64***	62
Became motivated to learn course content	68	73	67	78*	67	73	**65**	**70***	68

% agreement on 6 point scale (responses to three levels); * more than one standard error of measurement difference.

self-direct and become responsible for their own learning (62%). An important common student use was noted for this set of variables. Student who used the material in a collaborative manner reported greater affect development than did those who were involved in autonomous learning settings.

6 Conclusion

Overall, the students reported that the modules provided supplemental or alternative methods of instruction that could assist them in learning the core concepts presented in the introductory courses of Electrical Engineering. The majority of students also noted signs of individual outcomes pertaining to learning affect. Specifically, they perceived module usage as increasing their knowledge base and understanding in the specific content area, helping them to visualize problems, develop problem-solving skills, apply course content to new problems, and recall course content. Additionally, students were observed more active and engaged during module usage than during the lecture portion of the class.

Key variations were noted that support the acknowledgement of user characteristics and the need to ensure that the modules meet the needs of these users. Although both groups had overall positive perceptions of module use, male students had higher perceptions on several dependent constructs including the modules' relevance toward practice in the classroom, practice with real variables, their ability to tie content to new problems, and general problem solving. Male students also reported that the use of the modules better met their learning styles than did the female students and self-reported greater increases in knowledge and in confidence in their own ability. Female students, however, reported gains in knowledge on developing different patterns of problem solving than did male users.

Variations also were noted for perceptions based on whether learning took place in required autonomous or group settings. Those who did participate in group or collaborative use reported greater relevance to the course, greater gains in pictorial and visualizing skill, more direct recall, better application to new problem settings, and greater transference of learning. These students also reported more gains in knowledge, more motivation, confidence in mastery of material, greater interest in learning related topics, and greater gains in self-direction and control of learning than did those who worked alone. Students who worked alone found the module, itself, more useful and easier to understand, and reported that use improved their grades.

Students with no prior experience also reported positive outcomes. They perceived higher levels of transfer and problem solving skills than did those students with prior experience and reported greater gains of knowledge and confidence.

Overall, these data support the use of web-based learning modules for all types of learners but provide extra support for those that are constructivist in nature. Modules that require students to interact, develop their own cognitive framework, and present or share that knowledge with others appears to be especially useful. Further research is needed on the interaction of key student variables and how they moderate the impact of each for the five constructs under study. In addition, instructor/user characteristics also should be examined to determine if key modes of use might influence use. For instance, does the use of embedded module use impact outcomes

for selected types of students differently than does rehearsal or demonstration use? As designers, developers, evaluators, and users commit greater teaching roles to on-line or web-based tools answers to these questions must be found. The use of user centered approaches to module development and documentation will assist in this process and should be encouraged.

References

1. Lambropoulos, N.: User-Centered Design of Online Learning Communities. In: Lambropoulos, N., Zaphiris, P. (eds.) User-Centered Design of Online Learning Communities. Hershey, PA, IDEA Group, Inc. pp. 1–28 (2007)
2. Papastergiou, M.: Attitude as Predictor of Success in Online Learning. International Journal on E-Learning 5, 593–622 (2006)
3. Akhras, F.N., Self, J.A.: Modeling the Process, Not the Product, of Learning. In: Lajoie, S.P. (ed.) Computers as Cognitive Tools: No More Walls, vol. II, pp. 3–28. Lawrence Erlbaum Associates, Mahwah, NJ (2000)
4. Newman, D.L.: Taking Learner Characteristics into Consideration: Improving Development and Use of Technology Supported Learning. In: A paper presented at Human Computers Interaction (HCI) International Conference, Las Vegas, NV (2005)
5. Pierson, M.: NTLS VII: Key Research Questions for Technology and Core Content Disciplines. Journal of Computing in Teacher Education 22, 47–48 (2006)
6. Thompson, A.D.: Scientifically Based Research: Establishing a Research Agenda for the Technology in Teacher Education Community. Journal of Research on Technology in Education 37, 331–338 (2005)
7. Cheng, Y.C.: New Paradigm of Learning and Teaching in a Networked Environment: Implications for ICT Literacy. In: Wee Han, L.T., Subramaniam, R.(eds.) Handbook of Research on Literacy in Technology at the K-12 Level. Hershey, PA, IDEA Group, Inc. pp. 1–20 (2006)
8. Sugrue, B.: Cognitive Approaches to Web-Based Instruction. In: Lajoie, S.P. (ed.) Computers as Cognitive Tools: No More Walls, vol. II, pp. 133–162. Lawrence Erlbaum Associates, Mahwah, NJ (2000)
9. Law, E.L.C., Havannberg, E.T.: Quality Models of Online Learning Community Systems: Exploration, Evaluation, and Exploitation. In: Lambropoulos, N., Zaphiris, P. (eds.) User-Centered Design of Online Learning Communities. Hershey, PA, IDEA Group, Inc, pp. 71–100 (2007)
10. Nemeth, C.P.: Human factors methods for design: Making systems human-centered. CRC Press, Boca Raton, FL (2004)
11. Newman, D.L.: Rensselaer Polytechnic Institute Project Links: Mathematics and Its Applications Across the Curriculum. A technical report, Evaluation Consortium, University at Albany/SUNY (1998)
12. Newman, D.L., Passa, K.: The role of user characteristics in the development and evaluation of E-learning systems. In: Law, W.K.(ed.) Information resources management: Global challenges. Hershey, PA: Idea Group Publishing (in press)

Some Aspects of the Development of Low-Cost Augmented Reality Learning Environments as Examples for Future Interfaces in Technology Enhanced Learning

Alexander Nischelwitzer[1], Franz-Josef Lenz[1], Gig Searle[2], and Andreas Holzinger[2]

[1] University of Applied Sciences FH JOANNEUM, A-8020 Graz, Austria
School of Information Management
Digital Media Technologies Laboratory
alexander.nischelwitzer@fh-joanneum.at
[2] Medical University of Graz, A-8036 Graz, Austria
Institute for Medical Informatics, Statistics & Documentation (IMI),
Research Unit HCI4MED
andreas.holzinger@meduni-graz.at

Abstract. Augmented Reality (AR) applications are becoming increasingly available for everyday applications. In this paper, we concentrate on how to design and develop applications for educational purposes with the use of the ARToolkit and present an example from the area of human medicine. Since good Usability is essential, especially for non-expert end users, including both children and the elderly, we put particular emphasis on how to ensure good Usability. For both groups of end users, this technology has high potential and can be of great benefit. Within a small scale usability study amongst children, we could gain first experiences in their interaction with such applications. An essential factor is the notion of play, which could be beneficial to transport difficult learning material. On the basis of the gained insight, we provide an outlook for further work and we conclude that elderly could also benefit from this technology.

Keywords: Future Interfaces, e-Learning, Augmented Reality, Virtual Reality.

1 Introduction

The presentation of complex situations in ordinary books has always been a problem [1], [2]. Usually, the construction of a relevant mental model from a complex situation relies on the capacity of the learners to extract appropriate information from that situation [3]. Pictures, however, can support learners and, with computers, animated pictures can be manipulated for active exploratory learning, which allow the dynamic behavior of complex subject matter to be displayed [4].

However, in order to show detail in the third dimension, it is usually necessary either to *slice* the picture (for example anatomical images) or by using layers of schematic drawings.

Texts and images provide restricted possibilities to describe a difficult context, whereas three dimensional views of complex situations can sometimes increase

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 728–737, 2007.
© Springer-Verlag Berlin Heidelberg 2007

understanding [5]. By incorporating an interface that allows the reader to interact with the book and receive a feedback on the learning progress, everyone, independent of age, is offered a method that allows them to control their own learning progress and speed.

Although the application described in this paper was tested on school children between the ages of 7 and 13 years old, the software can be adapted to other age groups. Many older people are not cognoscente with modern technology [6], however this type of interactive learning is relatively easy to use and can be used to convey most subjects with relatively little assistance.

This paper describes an interactive Augmented Reality application for children, called My *Inside the body* Book (MIBB). The physical book (see figure 1) is used to interface reality and virtuality and presents the human alimentary system three dimensionally. In addition, the user can interact with the software and manipulate the image, adding and removing organs, by using the buttons mounted at the bottom of the book.

Fig. 1. Standard book compared to the Augmented Version (right)

2 Related Work

The Human Interface Technology Lab in New Zealand has been developing applications along the same lines. For example, by looking through a handheld device, the pictures in of sailing ships in the *Black Magic Book,* which tells the story of the America's Cup, become animated 3D models [7]. They have also developed an AR-Book called the *eyeMagic* book [8]. This project explores how a children's picture book can be transformed into a new type of reading experience by enhancing it with a three dimensional virtual content through the use of Augmented Reality.

Another comparable application to the MIBB is the Augmented Reality Volcano Kiosk, which is an exhibit that teaches people about volcanoes [9]. This also uses a handheld visor to visualize volcanoes, the ring of fire or the tectonic plates in three-dimensions (see figure 3). Interaction with the virtual objects is done through a slider.

An application that also uses the Augmented Reality Book approach is the *eyeMagic Book* [8] that allows the transition between reality, augmented reality and virtuality (see figure 4). This book can be read either with the use of a head mounted display [10], which makes virtual objects pop out of the pages of the book or just by means of the text and pictures in the book. The third possibility is to immerse oneself

into a total virtual world. The possibility of immersion, which provides end users with the experience of being surrounded by the environment has several opportunities for further research in psychology including object manipulation [11].

FIG. 2. AMERICAS CUP [7]

Fig. 3. Volcano [9]

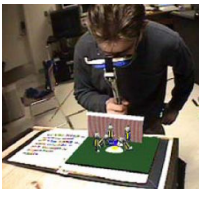


Fig. 4. eyeMagic book [8]

The MIBB has taken this a step further and, in comparison to standard applications [12], [13], has increased the interaction with the end user, which we consider to be essential in order to stimulate both motivation and engagement,. This has been achieved by implementing buttons within the MIBB. In addition, the user is asked to rotate the book and to examine the virtual objects from every perspective in order to get full benefit from the interaction possibilities of the story. This way, the MIBB itself is used as an interface between reality and virtuality. Several studies have provided evidence for the positive effects on learning through application of AR technologies [14], [15].

3 Design and Development of MIBB

The MIBB was developed with the Augmented Reality application software library ARToolKit [16], which permits the detection of black squares containing a graphical symbol within an image taken from a webcam. The position of the camera [user's eye] is then calculated with the help of those markers. In this way, the real world view can be merged with computer generated three-dimensional objects. The result of this *merger* is an Augmented Reality (AR). The process of detecting a marker and consequently extending the real world view with virtual elements includes the following steps: (1) take an input image of webcam; (2) convert the taken image to an binary image; (3) detect markers in binary image; (4) estimate pose and position; (5) draw the virtual image according to rotation and position of the marker.

Within an ARToolkit application, these five steps are executed continually. This loop gives the user an impression of an extended real environment. MIBB, which is made of cardboard, contains four pages. Instead of text and images, ARToolkit markers are printed on each of the four pages.

The markers, placed on the pages correspond with the three dimensional internal organs that currently appear to rise up from the page of the MIBB. Each page of the book is configured as a multi-marker page. This configuration avoids the occlusion of several markers without leading to an error drawing the internal organs. To use MIBB, the end user takes a seat in front of a projection screen. With the MIBB in the hand and a webcam directed onto the book, the augmented view becomes visible on the projection screen. It is also possible to use MIBB with a Head-Mounted Display

(HMD), which is particularly effective in displaying stereoscopic images [10]. This provides the users with a direct view onto the book and does not require them to switch between the book in their hand and the projection. The application was also successfully tested with an I-Window system (www.i-window.info), which projects from behind the screen.

4 Material Used as Content

MIBB was originally created as interactive story that guides the end user through the four pages of the book. Within the interactive story, three questions about the human alimentary system are answered: (1) What tasks does the digestion have? (2) Which organs take part in the process of digesting? (3) How do those organs work together to digest food?

The content of the book, intended for children aged from 7 to 13, is delivered by a female and a male voice. While the female voice provides the information about the different organs, the male voice poses questions and asks the end users to transport food and produce liquids needed for the process of digesting.

The MIBB starts when the end users take their place under the webcam and the marker on the cover of the book is tracked. After some basic information about the alimentary system, given by the female voice, the users are asked to open the first page, where they can see the gullet, the windpipe and the stomach. While getting general audio information about the displayed organs, the user is provided with the opportunity to superimpose comments onto the organs with the red button or to separate the organs with the grey button of the control pad.

To visualize further organs of the alimentary system on the following page of the book, the user has to answer questions concerning the organs that are currently displayed.

Therefore the male voice poses the question and gives three possible answers that are displayed beside the three buttons of the control pad in the book (see figure 6). Each of the answer possibilities is dyed in the colour of one of the buttons on the control pad. Pressing one of those buttons causes the female voice to confirm whether the answer was right or wrong. If the answer to the question was wrong, the user is asked to search for the correct answer by rotating the book and examining the virtual organs on the pages.

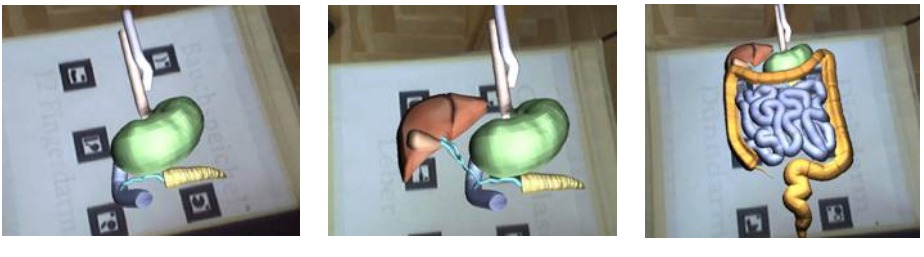


Fig. 5. Some pages of MIBB

Another restriction is that before the end user may move on to the following page of the book, the green button on the control pad must be used to virtually transport food through the gullet and stomach. On the following page further organs of the alimentary system are added to the organs that were displayed on the first page. After answering questions, transporting food and producing the liquids needed for digesting, the user can move on to the following page until the last page is reached, where the entire alimentary system is visible. Figure 5 shows some pages of the book and the virtual organs displayed on it.

5 Interaction

The physical book that the user holds in his hands can be used as an interface between the virtual world and the real world. Moving and rotating the book allows the user the chance to view the organs displayed on the page from every perspective. The examination of the organs through moving and rotating the book can help to correctly answer questions about the organs. To allow the interaction with the virtual organs on the four pages of the MIBB, an interaction control pad with three buttons was implemented at the bottom of the book (see figure 6). This plastic case contains a dismantled wireless Logitech Gamepad. The original buttons were removed and replaced with stain-less steel push buttons varnished in grey, red and green. To access the buttons of the Gamepad the Direct Input module of DirectX was used. When using the grey, red and green buttons, the following functions can be executed.

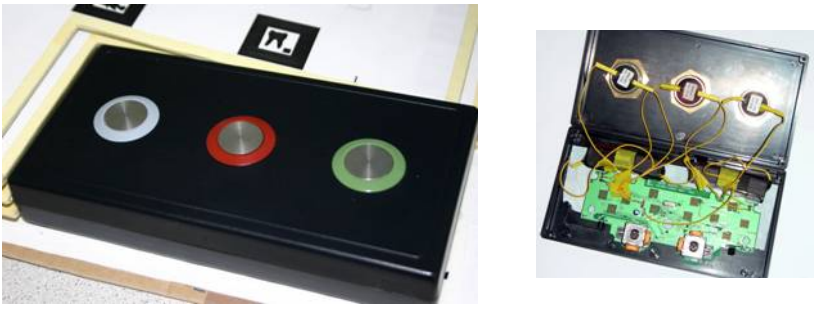


Fig. 6. The interaction control pad

By pressing the red button, the application displays an annotation to an organ. Repeated pressing shows another organ of the alimentary system annotated.

In addition, the organ is coloured yellow. Figure 7 shows an annotation for the pancreas. This function is especially necessary when the user wants to know the names of the organs in order to answer the questions in the interactive story.

This effect of displaying annotations is reached by replacing the current organ by a VRML-world, in which the organ is coloured yellow. The text and the arrow pointing to the organ are added with OpenGL library (see figure 7 oval).

The *grey button* allows the organs of the alimentary system (see figure 8) to be moved. This function leads to an improved understanding of the functionality of the

Fig. 7. Show annotations for organs

Fig. 8. Move organs

human body. Pressing the grey button again causes the organs to join together. All the organs displayed on the page of a book are loaded as separate VRML-worlds. This allows the execution of simple OpenGL transformations on each of the WRLs so it is possible to view each organ on its own (see figure 8 circle). As described in chapter 5, when the three possible answers are displayed at the bottom of the page, the user must press the button with the colour corresponding to the correct answer (see figure 9 circle).

Fig. 9. Answering questions

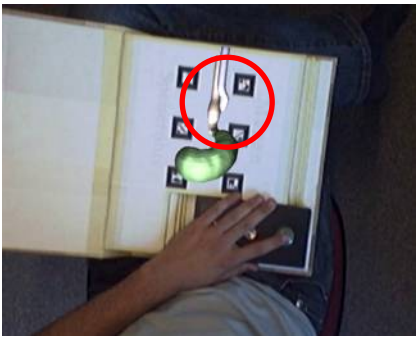


Fig. 10. Transport food through organs

To find the correct answer to the question, the virtual alimentary system on the page of the book can be examined from every perspective.

6 A First Usability Test

To gain preliminary experience of the implemented MIBB, a rapid usability test was performed [17]. The aim was, to find first answers to the following questions:

(1) Do children like using the MIBB? (2) Can children follow the instructions provided through the audio guides? (3) Is the interaction with the three buttons

Fig. 11. Some pictures taken during the Usability Testing (Book versus MIBB)

understandable? (4) Which parts of the MIBB can be improved? (5) Is learning with the MIBB easier than learning by reading a page of a children's book?

In order to find answers to these questions, N = 18 children between 7 and 13 years were invited to participate in a usability test. After requesting them to complete a questionnaire to discover the extent of their previous knowledge of the alimentary canal, one group, comprising of half the children, were asked to use MIBB while the other group was asked to read one page of an ordinary children's book about the alimentary canal.

Table 1. Questions, asked about the human alimentary canal during the usability test

Question	Answer 1	Answer 3	Answer 3
Which organ transports food to the stomach?	windpipe	gullet	duodenum
Which organ moves food to the duodenum?	liver	gall bladder	stomach
Where does the pancreas transports its liquids to?	colon	duodenum	gullet
How many lobes does the liver have?	2 lobes	4 lobes	1 lobe
Where does the gall bladder transports its liquids to?	liver	stomach	duodenum
Which organ transports food to the small intestine?	colon	duodenum	gall bladder
Where does the small intestine give the food to?	stomach	liver	colon
The appendix is a part of the …	small intestine	colon	liver

After playing the MIBB application and reading the page of the children's book, the test participants were again asked to fill in a questionnaire with the same questions as before. The following graphs show the result of the usability test (see figure 11 and figure 12).

Figure 12 shows the eight questions and the percentage of correct answers in each of the two questionnaires plus the answers selected within the MIBB application. Figure 13 shows the percentage of correctly answered questions in the questionnaire before and after reading a text about the alimentary system. When comparing those

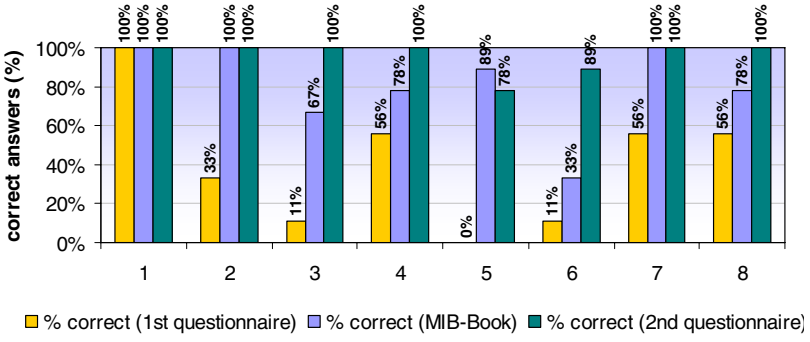


Fig. 12. Group A: Using the MIBB

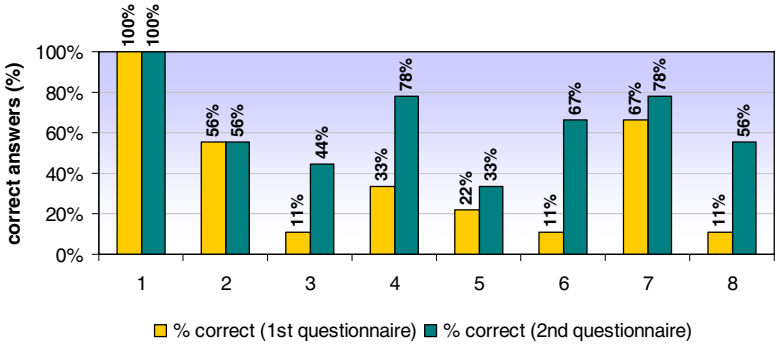


Fig. 13. Group B: Reading the corresponding text

two figures, it can be seen that the children who used the MIBB answered the questionnaire after the test better than those who just read the text (see green bars in figure 12 and figure 13).

Figure 14 shows the percentage of correctly answered questions within the two questionnaires plus the correctly answered questions within the MIBB application (blue bars). This data is compared to the correctly answered questions selected by test participants who read the text about the alimentary system.

The results of this first test show that children who used the MIBB acquired more knowledge about the alimentary canal than children who just read the text. This is possibly due to the three dimensional view of the virtual organs on the pages of the MIBB. In addition, the audio guides and the interaction of the end user with the virtual organs seem to improve the understanding of the functionality of the alimentary canal.

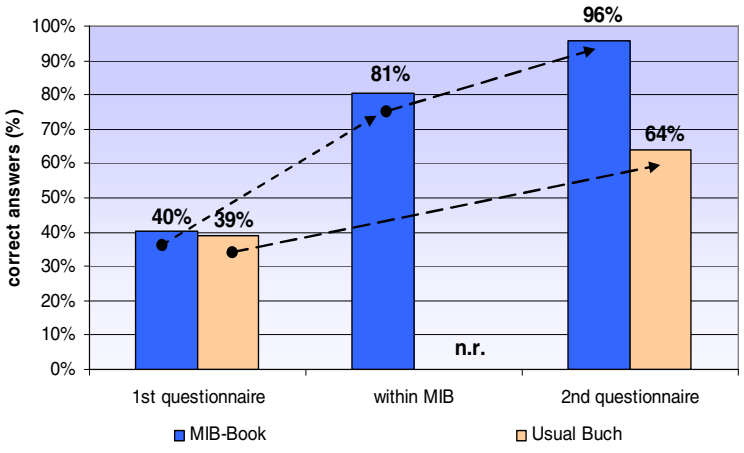


Fig. 14. Using MIBB

7 Conclusion and Future Work

This paper showed an innovative way of presenting information to children by using the *My Inside the body* Book (MIBB), which is a low cost AR application that displays the human alimentary system three dimensionally. MIBB allows the examination of some of the organs of the human body from every perspective. Consequently, the end users can interact with the virtual organs through buttons implemented at the bottom of the MIBB. A first preliminary test, performed with a group of children (N=18) between 7 and 13 years, showed the possible potential of AR in the area of learning. Combining the real world with three-dimensional, virtual content could also improve the understanding of complex subjects in other domains. However, more and detailed research will be necessary in order to evaluate possible positive effects, to generalize these findings and to incorporate the results into the development. Our next step will be to study this application with a larger population, and a wider range of ages, in particular older people. We are of the opinion that there are many possibilities to use this technology in other application domains in order to make complex things easier to understand.

References

1. Mandler, J.M., Parker, R.E.: Memory for Descriptive and Spatial Information in Complex Pictures. Journal of Experimental Psychology-Human Learning and Memor 2(1), 38–48 (1976)
2. Lamberts, K., Brockdorff, N., Heit, E.: Perceptual processes in matching and recognition of complex pictures. Journal of Experimental Psychology-Human Perception and Performance 28(5), 1176–1191 (2002)
3. Lowe, R.K.: Processing challenges of understanding complex animated pictures. Zeitschrift Für. Pädagogische Psychologie 12(2-3), 125–134 (1998)

4. Schnotz, W., Bockheler, J., Grzondziel, H.: Individual and co-operative learning with interactive animated pictures. European Journal of Psychology of Education 14(2), 245–265 (1999)
5. Guttormsen-Schär, S., Krueger, H.: Using New Learning Technologies with Multimedia. IEEE MultiMedia 7(3), 40–51 (2000)
6. Phang, C.W., Sutanto, J., Kankanhalli, A., Li, Y., Tan, B.C.Y., Teo, H.H.: Senior citizens' acceptance of information systems: A study in the context of e-government services. IEEE Transactions On. Engineering Management 53(4), 555–569 (2006)
7. Billinghurst, M.: Human Interface Technology (HIT) Lab (last access: 2006-08-29) http://www.hitlabnz.org
8. McKenzie, J.: The eyeMagic Book: A Report into Augmented Reality Storytelling in the Context of a Children's Workshop (last access: 2007-01-16) http://www.hitlabnz.org/route.php?r=prj-view&prj_id=18
9. Looser, J., Billinghurst, M., Cockburn, A.: Through the looking glass: the use of lenses as an interface tool for Augmented Reality interfaces. Computer graphics and interactive techniques in Australasia and South East Asia (2004)
10. Shibata, T.: Head mounted display. Displays 23(1-2), 57–64 (2002)
11. Loomis, J.M., Blascovich, J.J., Beall, A.C.: Immersive virtual environment technology as a basic research tool in psychology. Behavior Research Methods Instruments & Computers 31(4), 557–564 (1999)
12. Holzinger, A., Pichler, A., Almer, W., Maurer, H.: TRIANGLE: A Multi-Media test-bed for examining incidental learning, motivation and the Tamagotchi-Effect within a Game-Show like Computer Based Learning ModuleEducational Multimedia, Hypermedia and Telecommunication 2001. Association for the Advancement of Computing in Education, Charlottesville (VA) pp. 766–771 (2001)
13. Holzinger, A., Pichler, A., Maurer, H.: Multi Media e-Learning Software TRIANGLE Case-Study: Experimental Results and Lessons Learned. Journal of Universal Science and Technology of Learning 0(0), 61–92 (2006)
14. Heo, M., Chow, A.: The impact of computer augmented online learning and assessment tool. Educational Technology & Society, 8(1), 113–125 (2005)
15. Price, S., Rogers, Y.: Let's get physical: The learning benefits of interacting in digitally augmented physical spaces. Computers & Education 43(1-2), 137–151 (2004)
16. HITL: ARTOOLKIT Homepage (last access: 2007-01-15) http://www.hitl.washington.edu/artoolkit
17. Holzinger, A.: Usability Engineering for Software Developers. Communications of the ACM 48(1), 71–74 (2005)

Students' Attidudes Towards Novel Interfaces in E-Learning

Margit Pohl[1], Ilona Herbst[2], Franz Reichl[2], and Sylvia Wiltner[1]

[1] Institute of Design and Assessment of Technology, Vienna University of Technology, Favoritenstr. 9/E187, A-1040 Wien, Austria
{margit,sylvia}@igw.tuwien.ac.at
[2] E-Learning Centre, Vienna University of Technology, Gusshausstr. 28/E0152, A-1040 Wien, Austria
{reichl,herbst}@elearning.tuwien.ac.at

Abstract. There is still very little research about students' attitudes about e-learning. Such information seems to be necessary for a successful implementation of e-learning. Current research indicates that main advantages of e-learning are its flexibility and the provision of online learning material. A major drawback, as perceived by the students, is the lack of personal relationships. We conducted two focus groups to clarify these issues (N=54). The study supported the results from the literature. In addition, we found a few controversial issues as, e.g., preference for lectures/textbooks and tight schedules vs. preference for online learning and few deadlines. Another controversial issue was electronic assessment.

Keywords: students' attitudes, learner-centered design, academic teaching, university didactics.

1 Introduction

E-learning at universities becomes increasingly important. Nevertheless, it is sometimes not clear whether learners appreciate these novel forms of learning. Baker et al [1] and Sharpe et al [2] point out that there is still too little research about the attitudes of faculty and students about e-learning systems. Students are often enthusiastic and motivated by the new technical features of e-learning systems, but sometimes problems arise. When these problems are not fully understood and e-learning systems are introduced despite students' negative attitudes difficulties might come up. Collis and Moonen [3], for example, argue that students are not intrinsically motivated by the use of educational technology. When this technology is not integrated into the course-work, they will soon lose interest. As a consequence, the use of this technology will become less effective than traditional forms of teaching. More detailed information about students' attitudes about e-learning seems to be necessary. The following text describes an investigation which was conducted at the Vienna University of Technology in autumn 2006 in the course of the Delta 3 project. This investigation tried to identify major issues concerning the introduction of e-learning systems at the university. The method used for this investigation were focus groups. Major results from a literature review could be confirmed, and other important issues were raised in the focus groups.

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 738–747, 2007.
© Springer-Verlag Berlin Heidelberg 2007

2 E-Learning at Vienna University of Technology

Until 2005, e-learning at the Vienna University of Technology was provided on an experimental basis by projects initiated and carried out by a limited number of individual teachers. What was lacking was an institutional strategy for the university.

When the new university law which provided additional autonomy to the Austrian universities came into effect in 2004, the Vienna University of Technology founded an E-Learning Centre to:

- implement support services to teachers,
- provide easy-to-use e-learning tools to all teachers,
- network among the e-learning pioneers and other teachers,
- increase the number of lectures applying e-learning,
- develop a platform able to integrate the existing pioneering e-learning systems,

However, funding for this E-Learning Centre was limited. Such funding became available through the Delta 3 project (http://www.delta3.at/). With the Delta 3 project, 3 partner universities (Vienna University of Technology; University of Natural Resources and Applied Life Sciences, Vienna; Academy of Fine Arts, Vienna) focus their e-learning and e-teaching strategies on 3 target groups (teachers, learners in initial & continuing education, general public) through progress and synergies in 3 areas of competence (didactics, technology, design & usability). Internal and inter-university co-operation as well as trans-disciplinarity generate "public awareness of the science and arts" through relations between the corners of the 3 "e-learning triangles" as mentioned above. The project partners aim at a significant increase in the proportion of online-lectures, thus increasing the quality of the curricula by intensifying the learning process and by applying innovative didactic strategies for teaching complex content – with the objective of enabling otherwise disadvantaged groups of students to participate in courses. Gender mainstreaming as part of quality assurance ensures that e-learning will not discriminate against women or men.

Vienna University of Technology does not intend to replace face-to-face teaching entirely by e-learning, but it aims at enhancing the majority of lectures by electronic means. This takes place in the form of blended learning.

As mentioned above, one of the main goals of the Delta 3 project is the development of an integrative platform for the whole university. This platform had to fulfil the following criteria:

- open source,
- modular structure,
- potential to interact with existing e-learning and administrative systems,
- availability of tools for communication,
- usability,

Based on these criteria, the open source platform Moodle was chosen and adapted to the universities' requirements. The system is now called TUWEL (see Fig. 1).

At the Vienna University of Technology there also exists an electronic system for administrative purposes (TUWIS). In contrast to this system, TUWEL offers the

Fig. 1. TUWEL

possibility to upload complex material for lectures, to conduct tests, to develop WIKIs and to implement various forms of electronic communication (chat, forum, Weblogs etc.).

TUWEL conforms to the Vienna University of Technology's corporate design, and it offers an interface to existing administrative software systems at the university. In spring/summer 2006 a usability study was conducted which identified problems in the following areas: inconsistent location of buttons and inconsistent use of wording and terminology. During the second half of 2006, TUWEL was redesigned to remove the most serious of these problems. Several features have been implemented into TUWEL to enable activities required by teachers of Vienna University of Technology, e.g. group submissions of practical exercises and specific formats of tests.

TUWEL has been fairly successful and is increasingly used by teachers at Vienna University of Technology (see Fig. 2). This figure shows that the number of lectures TUWEL was used for was about 50 in the summer semester of 2006 and about 160 in the winter term of the same year.

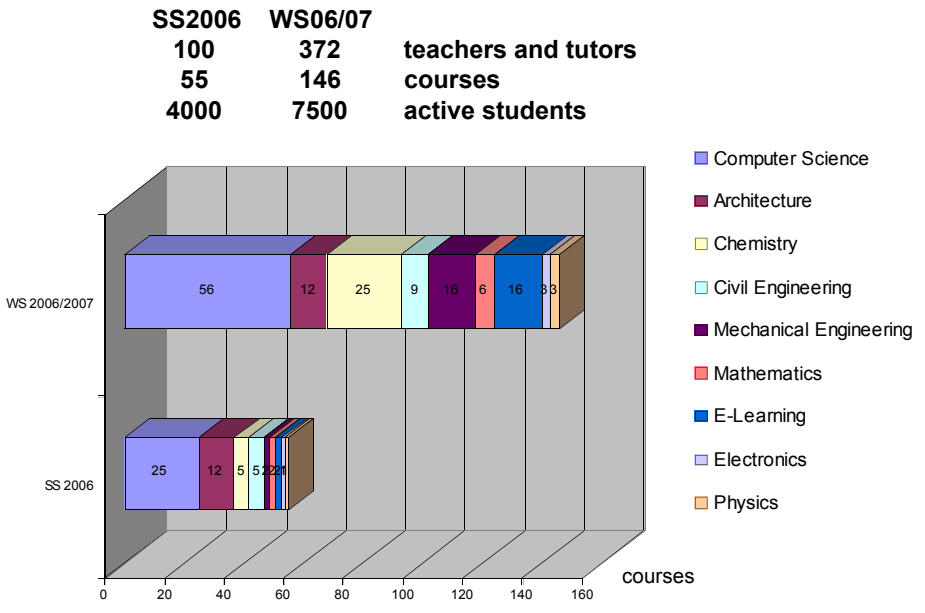


Fig. 2. Pattern of usage of TUWEL

3 Overview of the Literature

As mentioned above, there is still fairly little research about students' attitudes concerning e-learning. This is all the more surprising as many researchers in the area of e-learning adopt a learner-centered approach ([4], [5]). It is especially noteworthy that the learner experiences are in many cases different from those made by university staff, and that there are, furthermore, marked individual differences between individual students [2]. In addition, there is some indication that e-learning systems are not always used as intended by the designers and that learners often develop workarounds when they are not content with an e-learning system [6]. Therefore, it seems plausible to assume that speculation about students' attitudes and behavior will be problematic, and generalizations across students difficult. Detailed research in this area is necessary.

Although there is still too little research in this area, progress has been made in recent years. The study by Sharpe et al [2] is especially interesting in this context. The aim of this study was to get a comprehensive overview of experiences about blended learning in the UK. To reach this goal, the authors conducted, on the one hand, a literature search, and, on the other hand, an empirical study based on interviews in several universities. The authors state that student response to e-learning is mainly positive. Nevertheless, they find significant individual differences. In general, frequent usage of online resources can be observed. This conforms to our experience with TUWEL which shows that students also work during the night and on weekends.

One of the issues in e-learning seems to be that students get flexible access from home and from the campus. Students can study at their own pace and use their time more

effectively. This seems to be one of the most important advantages of e-learning for university students. Another significant advantage seems to be that students can access course notes at their convenience. In the literature study, the authors found that this is seen as more beneficial than online discussions and quizzes. It must be mentioned, however, that this raises the question of whether the availability of lecture notes and additional material does not have an unfavorable impact on attendance in lectures. The authors state that the reasons for attendance or absence in lectures is probably more complex and cannot only be explained by the availability of lecture notes.

Electronic communication is an essential part of e-learning. In principle, students are thrilled by this form of communication. Nevertheless, they do not want e-learning to replace face-to-face communication. They see e-learning and face-to-face teaching as complementary and both valuable. The use of electronic communication as such seems to be quite difficult because often students cannot be engaged to participate. Even when electronic communication is well integrated into a course, participation is still a problem. Another issue is online assessment and feedback. Online feedback seems to be very important and encouraging for students. Concerning online assessment there still seem to be some open questions. Individual differences may relate to nationality, gender, disabilities, learning style or other factors. These differences have to be taken into account although some of them seem to be more important than others.

The JISC conference [6] presents several different approaches to learners' experiences, among others the LEX study. This study is based on a qualitative approach (interviews plus artefacts such as learning diaries etc.). The target group is not only university students but also participants of continuing education courses and others. In this study differences between novices and immersive learners were found. Experienced e-learners were generally positive about e-learning whereas novice learners often lack confidence. Important aspects which influence students' attitudes to e-learning are, e.g., possibility of distraction (which might be bigger at home) or the opportunity to work individually at one's own pace. Some students mentioned that computer-based assessments were a big advantage for them. Some students use communication software other than that provided by the e-learning platform (e.g. MySpace or MSN messenger). This might be a problem in some cases.

Tait [7] describes an older study in which the replacement of face-to-face lectures by rather comprehensive written material with pictures and videos presented on the computer is analysed. In principle, the attitude to e-learning in this form is fairly positive. Students mentioned one important advantage of e-learning, that is 'learning at one's own rate'. This is related to a second, less important advantage: 'fit it into your timetable'. On the other hand, no single drawback was identified by a majority of the students. Some students mentioned that e-learning was boring which might be related to reading a large amount of text on the screen. A considerable minority also stated that they felt the lack of communication with other people (lecturers, other students) negative.

Baker et al [1] report the highlights of a comprehensive study of almost 300 higher education institutions in the USA. In this study students as well as faculty were interviewed. An important result was that the target group thinks that old-fashioned personal communication is still more effective than advanced technology. The authors of the

study argue that this might change with the implementation of more tools for electronic communication. Another result is that faculty and students' attitudes about e-learning highly depend on the amount of support they get from their institutions. For e-learning to work smoothly, regular communication between students and teachers is necessary.

Mason [8] describes an investigation made at the British Open University. She also reports general enthusiasm, but she also points out that there are several drawbacks. Paradoxically, she states that new media make learning less flexible. When learning with the materials offered by the Open University students have to use several different media in a prescribed way (tape, programming environment, conferencing system etc.). This sometimes makes it difficult for students to fit learning into their time-schedules. Another drawback is that learning communities are welcomed by a certain group of learners but refused by others (especially those with a very tight schedule who do not benefit from extended communication with peers). In addition, students tend to resist the substitution of face-to-face meetings by computer conferencing.

A report by the Learning Technology Support Service [9] of the University of Bristol relates several results concerning students' attitudes about e-learning. Again, students generally find e-learning positive, but are sceptical about aspects they had not experienced before. The most important features of e-learning are the Web (to look for supplementary resources), distribution of learning materials via the Internet and getting administrative information. Key benefits are more practice, learning at one's own pace and greater access to information, main concerns are the potential reduction of face-to-face contact and that learning might become more mechanical.

Motschnig-Pitrik [10] and Motschnig-Pitrik and Holzinger [11] experimented with new pedagogical approaches combined with e-learning. She reports that the students' motivation was very high, and that they felt they had had a very intensive learning experience in contrast to the rather superficial learning in traditional lectures.

The research described above indicates that students value the flexibility of e-learning as one of its major advantages. The lack of face-to-face communication is sometimes mentioned as a problem. Innovative forms of e-learning are seldom used and not always appreciated although if designed well are seen as advantageous. Electronic communication is sometimes seen as a problem although, if designed well, this also can be one of the major advantages.

4 Description of the Study

Students are one of the target groups of the Delta 3 project. Their positive attitude to e-learning is an essential condition for the success of this form of learning. Therefore, we decided to conduct a study to get more detailed information about this topic. As a method of investigation we used focus groups [4]. There are several reasons for this decision. One is that because of the lack of very systematic previous research a more open approach seemed to be advisable. In this way, we also wanted to get information about topics which were not anticipated. Based on the available literature and personal discussions with students, we also assumed that there are several controversial questions concerning e-learning (e.g. attendance at lectures, usage of electronic communication, electronic assessment). We thought it advisable to discuss these questions in groups to

get more pronounced opinions. In addition, focus groups are a possibility to get a broad range of opinions and a good overview of a topic in a relatively short time. For a study aimed at getting first basic information about an area of interest focus groups, therefore, seemed to be advantageous. The only problem was that the groups as such were fairly large (approx. 30 persons per group). At first, we thought that this might cause serious difficulties but in the end we realized that the group size was manageable.

The subjects of the investigation were undergraduate students of Computer Science. They all took part in a lecture called "Networked Learning". Most of these students are experts in using computers and have no negative attitudes about computers. It can also be assumed that a majority of students is interested in the topic of e-learning. All the students had some basic experience in using e-learning platforms as there is a fairly complex electronic administrative system used by the Vienna University of Technology (TUWIS). This systems enables teachers to give information about lectures to students, to load up lecture notes, to mark assignments and to send messages to students. Students can register for lectures or tests in this system and they can take part in electronic discussions. Some of the students also had experience with TUWEL (which is connected to TUWIS but offers a broader range of features). Before the focus groups students got an introduction into this methodology and some information about the topics to be discussed. There were two focus groups. The total number of subjects was 54 (focus group 1 29, focus group 2 25). The single focus groups lasted about 90 minutes. One of the co-authors acted as moderator for the focus groups, the other three wrote protocols.

The focus groups were moderated according to a discussion guide. This guide was partly based on the literature described above and partly on literature about the design of e-learning systems ([12], [13], [14]). The topics of this guide were:

1. the students' opinion about the main problems and the main advantages of e-learning systems,
2. electronic communication and lack of face-to-face contact,
3. motivation (how to motivate students when learning is more open and flexible),
4. the design of learning material and lecture notes (use of pictures and videos, use of practical examples, hypertext, etc.),
5. goals of studying (personal development vs. preparation for a specific job)
6. electronic assessment,

The first question had to be answered by everyone. As expected, the students also discussed the other questions when they answered the first question (e.g., they said that one of the major disadvantages of e-learning was the lack of face-to-face communication). The other questions were answered in a free discussion. In the free discussion, about half of the students in both focus groups took part. The discussions were very lively, and the students were specifically asked to make critical remarks and not to hesitate to utter their personal opinions.

5 Results

A study based on focus groups cannot give representative results but only a first glimpse of relevant problems. Nevertheless, we did notice that some of the issues were mentioned

again and again. Greater flexibility of e-learning as an advantage was mentioned very often, and the lack of personal communication with other students or teachers was a very common argument against e-learning. This conforms to the results in the literature described above. In addition, a great variety of advantages and disadvantages of e-learning were discussed. The possibility to download lecture notes was mentioned fairly often. Some students also said that e-learning enabled them to access more information. It must be mentioned, however, that a few students found face-to-face lectures better. Some students remarked that taking notes in lectures forced them to think about the material presented there. Students also were aware of the problem that downloadable lecture notes motivated them less to go to lectures which might be a problem. Several students found e-learning platforms confusing. A problem which was also mentioned was that most teachers only used the Internet for uploading traditional lecture notes. They did not use it for innovative e-learning programs with interactive features. Some students also remarked that interactive examples would be very valuable to get more practice. Students also wanted more communication with other students or teachers. A problem mentioned by some students was privacy. They felt more observed by lecturers. TUWEL, in principle, allows lecturers to get information about usage patterns of students, although right now only aggregate data is provided to lecturers.

Electronic communication is very attractive for the students who took part in the study, but they use it only selectively. As mentioned above, personal contact is, in some situations, still important for them. To develop friendships or to coordinate complex group work, students still prefer to meet personally. One student mentioned that it is difficult to represent complex issues in the Internet. This is still easier with a pen and a piece of paper. Students also pointed out that too many participants make communication difficult. On the other hand, they do exchange a lot of information via Skype or instant messaging, for example, before tests. This information is apparently not so complex and related to the test for which they have to learn. One student also mentioned that in an electronic forum people only post answers when they are 100% sure. This is different in face-to-face communication.

Motivation was also a relevant issue. This issue was very controversial. Some of the students argued that they need the time schedule offered by regular lectures to remain motivated. Other students argued against that. They think that they are disciplined enough and do not need lectures. Some students suggested that in e-learning, deadlines for the submission of assignments are possible. In this way, motivation might not be such a problem. One student remarked that such deadlines are in conflict with the flexibility of e-learning (seen by many students as an advantage).

The design of learning material was not such an important issue. In this context, students predominantly discussed the problem that it is only possible to access information from lectures for which one is registered (This is also due to copyright reasons). This is contrary to the intentions of the Internet and completely different to the huge network of information offered by the World Wide Web.

There was also a controversy about online assessment. Students were rather critical. They think that online assessment might be viable for multiple choice tests or mathematics tests. Some students mentioned that they are nervous during tests which makes

typing difficult. They argued that automatic assessment makes typing errors problematic whereas lecturers might be more tolerant of such mistakes.

The topic 'goals of studying' was not discussed in both focus groups because of time restrictions.

6 Conclusion

Literature indicates that there is still too little research about students' attitudes about e-learning. The study described in this text tries to give a first overview about this topic. The aim of this study is to inform the implementation process of the e-learning platform of the Vienna University of Technology and, thereby, make design decision easier and more efficient. The study is not representative for all students at this university. It must be mentioned, however, that Computer Science students, who formed the sample in this study, make up a high percentage of students at our university.

Some of the results conform to what can be learned from the literature. Students think that main advantages of e-learning are an increase in flexibility and the possibility to access relevant information. An important disadvantage is the perceived loss of personal communication. Apart from that, students have differentiated views about e-learning. An example would be that some prefer tight schedules and a higher degree of face-to-face learning, others would prefer e-learning with few time constraints. Online assessment is also a controversial issue.

Students' attitudes about e-learning seem to be a relevant and complex topic. We intend to conduct further, more systematic research in this area using questionnaires and interviews.

Acknowledgements. The Delta 3 project is supported by bm:bwk in the framework of the programme e-learning/e-teaching Strategies at Universities and Fachhochschulen.

References

1. Baker, M., Boggs, R., Arabasz, P.: Student and faculty perspectives on e-learning support. Educause Research Bulletin. 3(16) (2003)
2. Sharpe, R., Benfield, G., Roberts, G., Francis, R.: The undergraduate experience of blended e-learning: a review of UK literature and practice. Technical report, The Higher Education Academy (October 2006) (last seen 31.12.2006) http://www.heacademy.ac.uk/research/Sharpe_Benfield_Roberts_Francis.pdf
3. Collis, B., Moonen, J., Collis, B., Moonen, J.: Flexible Learning in a Digital World. Kogan Page, London (2002)
4. Quintana, C., Carra, A., Krajcik, J., Soloway, E.: Learner-centered design: Reflections and new directions. In: Carroll, J.M. (ed.) Human-Computer Interaction in the New Millenium, pp. 605–626. Addison-Wesley, Boston, San Francisco, New York (2002)
5. Holzinger, A., Motschnig-Pitrik, R.: Considering the human in multimedia: Learner-centered design (lcd) & person-centered e-learning (pcel). In: Mittermeir, R. (ed.) Innovative Concepts for Teaching Informatics. Carl Ueberreuter, Wien, pp. 102–112 (2005)
6. Minshull, G., Mole, J. (eds.): JISC – Learner Experiences of e-Learning. In: The proceedings of the JISC Online Conference: Innovating e-Learning 2006, Cheltenham, Direct Learn Services Ltd (2006)

7. Tait, K.: Replacing lectures with multimedia CBL: Student attitudes and reactions. Instructional Science 26, 409–438 (1998)
8. Mason, R.: Globalising Education. Trends and Applications. Routledge, London, New York (1999)
9. Learning Technology Support Service: E-learning and virtual learning environments. What do staff and students think? Highlights from recent surveys on the use of virtual learning environments and e-learning at the University of Bristol (2005) (last seen 01.01.2007) http://www.ltss.bris.ac.uk/publications/reports/eLVLE%20survey.pdf
10. Motschnig-Pitrik, R.: Person centered e-learning in a major academic course: What are the results and what can we learn from them? In: Proceedings of the 4th International Conference on Networked Learning (NLC), Lancaster, UK (2004)
11. Motschnig-Pitrik, R., Holzinger, A.: Student-centered teaching meets new media: Concept and case study. IEEE Educational Technology & Society 5(4), 160–172 (2002)
12. Clark, R., Mayer, R.: E-Learning and the Science of Instruction. John Wiley, San Francisco (2003)
13. Pohl, M.: Guidelines for the design and evaluation of e-learning systems. In: Proceedings of the EDEN '06 annual conference, Vienna, Austria pp. 542–547 (June 2006)
14. Holzinger, A.: Multimedia Basics,Learning, vol. 2. Laxmi-Publications, New Delhi (2002)

The Impact of m-Learning in School Contexts: An "Inclusive" Perspective

Francesca Pozzi

Institute for Educational Technology - National Council of Research
Via De Marini, 6 16149 Genoa, Italy
pozzi@itd.cnr.it

Abstract. M-learning takes advantage of the use of mobile technologies (i.e. mobile phones, personal digital assistants, handheld computers) and allows students to carry out activities in a variety of settings and according to different paces. The paper discusses the main changes and challenges brought about by the use of m-technologies in school contexts with a special attention on one hand to the limitations posed, and on the other to the possibilities for enhancing students' inclusion.

Keywords: mobile, learning, impact, school, inclusion, technologies.

1 Introduction

Even if the use of ICT in school settings has become quite common, in most cases technology is something that is not really integrated in the learning activity, but rather it is a supplemental resource that is occasionally exploited.

This is due to a series of reasons, including the fact that the technological devices are often placed within laboratories and, in order to carry out ICT-based activities, students and teachers have to move from their classroom to the lab, thus causing organizational problems, as well as an overall disruption in the learning activity itself.

Taking advantage of the use of mobile technologies, m-learning should in principle allow students to carry out ICT-based learning activities in a variety of settings, ranging from the traditional places where learning usually occurs (classroom, lab, home, etc.), to places that are even far away from the "typical" ones. Furthermore, m-technologies should support the possibility for students to learn at their own pace. From this perspective, m-learning should offer a lot of advantages and allow the learning event to become really "time and space independent" [1]. In addition, mobile technologies support collaboration and information sharing, thus allowing group working, both face-to-face and at a distance.

All these features make mobile learning a very flexible approach that seems to be appropriate for addressing different needs and achieving different learning objectives.

Nonetheless, it is evident that the availability of m-devices in schools does not guarantee per se the learning experience to become a "m-learning" event; on the opposite: in order to effectively integrate m-devices in the pedagogical plan, the overall activity needs to be redesigned, taking into account that the very nature of m-devices affect several aspects of the teaching and learning process.

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 748–755, 2007.
© Springer-Verlag Berlin Heidelberg 2007

The paper provides an overview of the main changes and challenges brought about by the use of m-devices in school contexts and focuses on one hand on their ability to "include" students in the learning activity, and on the other on the limitations they pose.

2 Defining m-Learning

The very nature of "learning" is in itself "mobile". Vavoula and Sharples [2] state that: "learning is mobile in terms of space, i.e. it happens at the workplace, at home, and at places of leisure; it is mobile between different areas of life, i.e. it may relate to work demands, self-improvement, or leisure; and it is mobile with respect to time, i.e. it happens at different times during the day, on working days or on weekends".

From this perspective, one could argue that mobile technologies simply provide an "extension" to the "natural mobility of learning".

When looking at the literature for a definition of "m-learning", it is possible to find "technology oriented" definitions, as for example the one provided by Hoppe et al. [3] who state that "m-learning is e-learning using mobile device and wireless transmission". In such definitions the focus is on the kind of technology used which, from time to time, may include: palmtops (or handhelds), Personal Digital Assistants (PDAs), tablet PCs, mobile or smart pones, special handhelds (e.g. e-books, dictionaries, graphical calculators, spellcheckers). It is to be noted that, generally speaking, laptops – which in principle are "mobile devices" – are not included in such definitions, in that their portability is limited and their size and weight do not allow an *everywhere* use.

Other definitions of m-learning are more "learner-centered": O'Malley et al. [4] define m-learning as: "any sort of learning that happens when the learner is not at a fixed, predetermined location, or learning that happens when the learner takes advantage of the learning opportunities offered by mobile technologies."

Some other definitions stress the relationship between e-learning and m-learning. Petrova [5] defines mobile learning as "a form of e-learning" to be distinguished from "online learning". Petrova states that while the "mobile learner" is typically one who has got his/ her private device and who consequently can access the device whenever and wherever s/ he wants, the "online learner" is one who typically shares a desktop computer with others, having access to it only at fixed and pre-determined periods of time.

In the present paper we refer to "m-learning" as the *social*, *contextualized* and *individualized* learning process that takes advantage of the use of mobile devices in school settings.

3 M-Learning and Students' Inclusion

A lot of studies agree that m-learning can contribute to really integrate technology in the classroom, as opposed to the present occasional, supplemental use of technology.

Klopfer et al [6] argue that five characteristics of mobile devices allow them to produce unique educational affordances; such characteristics are: *portability*, that

allows to take the devices everywhere; *social interactivity*, that supports data sharing and collaboration among face-to-face students; *context sensitivity*, that is the possibility to gather and respond to real or simulated data in a specific context; *connectivity*, i.e. the availability of the network; *individuality*, that is the possibility to personalize activities on the basis of the individual needs.

According to [7], thanks to their ability to "engage students in creative, collaborative, critical and communicative learning activities", m-technologies "offer new solutions to traditionally problematic contexts of information delivery… The number of learners whose needs may be met by m-learning includes mature-aged, gifted, international and remote learners, as well as those with cognitive, behavioral, or social problems, or with physical of mental difficulties [8; 9; 10]. Mobile devices can help improve literacy and numeracy skills; encourage independent and collaborative learning experiences; identify areas where learners need assistance and support; mitigate resistance using ICTs; engage reluctant learners; enable learners to remain more focused for longer periods and promote self-esteem and self-confidence [11]" [7].

The impact of mobile technology on education could potentially be even stronger, because the constant exposure of students to digital technologies makes them "native" digital literates and thus the adoption of mobile learning should turn out to be particularly motivating and "natural" for them [7].

Despite this widespread enthusiasm, it is a fact that m-learning has not yet seriously impacted on educational contexts and the projects addressing the adoption of m-technology in school settings can still be regarded as spearheads [12]. This is perhaps due to a series of reasons of different nature (organizational, economical, technological, methodological) and it is beyond the scope of the present paper to analyze them. Certainly, it is worthwhile to further explore the field, since it seems that the spectrum of possibilities offered by this new generation of technologies is in some way even wider than that "announced" by the introduction of ICT in schools. As it will be illustrated in the following, such range of possibilities seems to open new horizons as far as students' inclusion is concerned. Despite the limitations they pose, it seems that today mobile devices can really open the school to a process of innovation, "finally allowing the impact of technology to match its promise" [13].

3.1 M-Learning and Inclusion: Main Challenges

Assuming that a teacher may dispose of mobile devices for all the students, s/he should wonder why to use them, for which educational purposes.

Starting from the definition we have provided of mobile learning, we could infer that the use of handhelds is worthwhile when the teacher aims at stressing:

- the *social dimension* of learning, by enhancing communication and collaboration among peers
- the *contextualization* of learning, by providing possibilities to "work on the field"
- the *individual dimension* of learning, by personalizing the activities according to the individual needs.

Gay et al. [14] provide an interesting categorization of the learning objectives that may motivate the use of mobile technology, which is shown in the following table:

Table 1. Objectives of m-technology in education [14]

Level 1	*Level 2*	*Level 3*	*Level 4*
Productivity	Flexible physical access	Capturing and integrating data	Communication and collaboration
Sample applications			
Calendars Schedule Contact Grading	Local database Interactive prompting Just-in-time instructions	Network database Data collection Data synthesis Mobile library	Real-time chat Annotations Data sharing Wireless e-mail
Content-intensive Users: individual Mostly asynchronous Information storage Hardware-centered Isolation			Communication-intensive Users: group Mostly synchronous Knowledge construction Network-centered Interconnection

The table enlightens that mobile technologies may address a certain number of objectives, ranging from "individual-centered" (left side of the table) to "group-centered" (right side) ones. Both the ends of the continuum may be regarded as promoting students' inclusion.

Inclusion is seen as a process of addressing and responding to the diversity of needs of all learners, comprising those with "special needs", i.e. those with intellectual or physical disabilities and learning impairments, those who are disaffected, "hard to reach" or even those who cannot attend school for personal or family/cultural reasons.

As we have already stated, one of the most interesting aspect of m-learning is the possibility to "personalize" the activities according to the individual needs of students. This may imply for example using functionalities, such as the calendar or an the automatic reminder, to support those students who have difficulties in scheduling their work.

Mobile devices may also provide the possibility to a student who cannot attend the class with regularity, to participate to the lessons and/or classroom events by using his/her personal handheld.

In other cases, the personalization may imply allowing students with cognitive difficulties to learn and work at their own pace, by providing them with specific activities or supplemental materials to be used according to their needs.

Or personalization may simply refer to the involvement in collaborative activities of students who usually work in isolation, so that they can use their palmtop for communicating, discussing and sharing materials with others.

Thus, mobile learning can in principle reach a variety of learners and may help the teacher in the difficult task of involving as many of them as possible.

Unfortunately, nowadays there are still limitations to face when adopting such technologies in educational contexts. Such limitations are first of all technical problems still to be effectively solved.

For example, one of the most important limitations related to the use of mobile devices in classroom, is the shortage of materials available for "mobile use": the contents very often need to be designed *ad hoc* or re-adapted from the existing ones in order to make them readable from the small screen displays and both these processes seem to require a certain effort. At the moment the teacher is usually forced to either produce the material by him/herself, or select the appropriate sources to provide to students.

Besides, if the teacher wants to adopt mobile technologies for inclusion purposes, the full accessibility and high-quality interaction of devices, applications and services should be guaranteed.

On the contrary, there are "objective" barriers created above all by the small size of PDAs in particular and portable devices in general, that make them not fully accessible.

First of all, the small size of the body may prevent the use to people with manual dexterity problems, who may find it difficult to hold the PDA while working on it, or to use the stylus.

But the small size may represent a problem for blind and visual impaired as well, who may suffer from the size of the screen, its low resolution and poor lightning and from the low tactile and visual quality of both screens and buttons.

Furthermore, people with speech and language difficulties and those with learning and /or cognitive difficulties may find it difficult to understand the graphical interfaces, the technical language that is often used in the manuals and technical documentation in general, and may suffer from a lack of visual representations (e.g. icons) of actions or commands.

Even if designers of mobile devices are always looking for better and more usable solutions, it is a fact that, at the moment, before considering the possibility to adopt such technologies, the teacher should take into account such constraints.

3.2 M-Learning and Inclusion: Main Changes

A mobile learning activity may be carried out in the classroom or outside the school and the teacher may choose to let the whole class work on the same task or divide it into groups, according to the needs.

Such a scenario is not new at all in the educational panorama and could happen independently from the use of technologies, being them mobile or not.

But let's suppose each student being equipped with an handheld of his/ her own, which is on his/ her desk together with books, pencils and paper. In this new scenario, it is possible for the teacher to launch any ICT-based activity directly in the classroom and whenever this is necessary, because this won't cause any disruption in the activity itself. The class won't need to transfer into the computer laboratory and each student (and/or each group) will be able to use ICT if and when she needs it.

Besides, another innovative advantage brought about by the integration of mobile devices in the school practice, consists of the possibility to adopt a variety of learning approaches.

Naismith et al. [15] have elaborated an activity-based categorization of mobile technologies and learning, where "mobile" activities are associated with their underpinning learning theories. According to their conceptualization, one may adopt

behaviorist approaches for example by setting up drill and practice exercises to be delivered through mobile devices; as a matter of fact instant messaging (e.g. SMS) can be fruitfully used for delivering tests and questionnaires.

On the other hand, more constructivist and situated approaches can be supported by handheld computer, for example by making students gather and elaborate data on the field during experiments.

Furthermore, socio-constructivist approaches may be enhanced by exploiting the ability of mobile devices to provide communication and collaboration services, to be used both during face-to-face sessions and at a distance.

Even if one could argue that such activities and approaches were already possible using stable workstations, the real, new potentiality afforded by mobile technology is the possibility to easily switch from one approach to another without disrupting the activity, i.e. shifting from an individual activity to a collaborative one, from a frontal presentation to a working session, on the basis of the learning objectives and on the students' needs. The frequent and easy switch between the virtual and the real space is also important from a cognitive point of view, above all in an age where acquiring knowledge means above all thinking critically and solving complex problems by retrieving information in context and using multiple sets of facts [16].

From an "inclusive" perspective, this means that during the learning activity, each student (or group of students) may indifferently be "virtually" or "physically" present, working on a personal or on a group assignment, and above all s/he may change this status softly and dynamically.

From this perspective, mobile devices are not at all a *surrogate* of desktop or laptop PCs, i.e. they can not be viewed as simply providing more portable versions of the learning activities that are currently supported on more static machines [15]. Thanks to mobile devices, the technology may become an integrated resource in the learning environment being "transparent" to the student and allowing the teacher to fully exploit its potentiality.

The teacher should be aware that the ownership of mobile devices by students provide them with a great potentiality, because in principle this allows them to become active constructors of their learning process. As a matter of fact, the introduction of mobile learning in the school practice should enhance a more autonomous learner's role [17], where students are allowed to make suggestions and have a certain degree of freedom in choosing when and how they use their device.

While the teacher is in charge of providing the general outline of the activity, it would be up to the single student to "shape" the activity, according to his/her needs, pace and interests. In this scenario, the teacher becomes a mentor, able to support the learning process by catching the stimuli provided by the students and directing them towards the learning objectives.

This implies changes in the competences required by the teacher, who is not any more the "knowledge repository" who "pours" the contents into the students' minds, but rather a facilitator of learning processes, the one who orchestrates all the resources and manages the dynamics towards the achievement of the learning objectives.

Besides, at a deeper analysis of the new panorama we have sketched, it soon becomes evident that great changes are required not only to the single teacher, but rather to the overall school system as well.

As a matter of fact, the school should guarantee a flexible organization, so to allow the teacher to fully exploit mobile learning potentialities.

Moreover, the school should provide support to the teacher who adopts mobile devices in his/her practice, not only in terms of proving the suitable equipment and infrastructure, but also affording a constant technical support.

The new teacher's role and students' attitude should also find a counterpart in the curriculum and the evaluation model, that should demonstrate a certain flexibility, so to allow the teacher to take into the due consideration the stimuli provided by the learners.

If in principle mobile learning presents characteristics that indicate it as a possible solution for tackling students' inclusion in class, on the other hand, great changes are required to the school system in order to integrate mobile devices as effective learning resources.

4 Conclusions

The present paper focuses on the use of mobile technology in school settings.

When looking at the literature and analyzing the experience in the field, it is easy to realize that - despite the great enthusiasm and the optimism arisen by such technology - the impact on the educational system is still quite low and the experiences carried out are still to be considered as spearheads.

The case studies are even less, if we consider those addressing people with special needs and this is because of the several limitations, mainly technical, that still mobile devices provide to the end user.

Still, there is accordance among the authors on the idea that mobile learning can help in engaging all the students and even more, those with difficulties.

Certainly, mobile technologies have features and characteristics of their own, which make them different from ICT in general, and thus a particular attention has to be devoted to them and *ad hoc* reflections are to be carried out. In particular, inclusion could be effectively addressed with a massive adoption of mobile learning, due to the fact that such technologies may allow a more individualized learning process.

In our view, there is still a lot of work to do in the field and new, challenging research questions are raising. As it has been up to now for the introduction of ICT in education, even the use of mobile technologies in school contexts may act as a propulsive agent of innovation, thus enhancing the school to face the new needs of the society and finally move it towards inclusion.

References

1. Roschelle, J., Pea, R.: A walk on the WILD side: How wireless handhelds may change computer-supported collaborative learning. International Journal of Cognition and Technology 1(1), 145–168 (2002)
2. Vavoula, G.N., Sharples, M.: KLeOS: A personal, mobile, knowledge and learning organization system. In: Milrad, M., Hoppe, U., Kinshuk (eds.) Proceedings of the IEEE International Workshop on Mobile and Wireless Technologies in Education (WMTE 2002) (Aug 29-30) Vaxjo, Sweeden pp. 152–156 (2002)

3. Hoppe, H.U., Joiner, R., Millard, M., Sharples, M.: Guest editorial: wireless and mobile technologies in education. Journal of Computer Assisted Learning 19, 255–259 (2003)
4. O'Malley, C., Vavoula, G., Glew, J.P., Taylor, J., Sharples, M., Lefrere, P.: Guidelines for learning/teaching/tutoring in a mobile environment. MOBIlearn deliverable D4.1 (2003) (accessed on February 10, 2007) http://www.mobilearn.org/download/results/guidelines.pdf
5. Petrova, K.: Mobile learning using SMS: A mobile business application (2005) (accessed on February 10, 2007) http://www.naccq.ac.nz/conference05/proceedings_04/petrova_mlearn.pdf
6. Klopfer, E., Squire, K., Jenkins, H.: Environmental Detectives: PDAs as a window into a virtual simulated world. In: Proceedings of IEEE International Workshop on Wireless and Mobile Technologies in Education, Vaxjo, Sweden, pp. 95–98. IEEE Computer Society, Washington, DC, USA (2002)
7. Cobcroft, R., Towers, S., Smith, J., Bruns, A.: Mobile learning in review: Opportunities and challenges for learners, teachers, and institutions. In: Paper presented at Learning on the Move, Brisbane, Australia (2006) (accessed on February 10, 2007) https://olt.qut.edu.au/udf/OLT2006/gen/static/papers/Cobcroft_OLT2006_paper.pdf
8. Savill-Smith, C., Kent, P., The use of palmtop computers for learning: A review of the literature. London, UK: Learning and Skills Development Agency (2003) (accessed on February 10, 2007) http://www.m-learning.org/docs/the_use_of_palmtop_computers_for_learning_sept03.pdf
9. Strom, P.S., Strom, R.D.: Personal digital assistants and pagers: a model for parent collaboration in school discipline. Journal of Family Studies 8(2), 226–238 (2002)
10. Rodríguez, P., Nussbaum, M., Zurita, G., Rosas, R., Lagos, F.: Personal digital assistants in the classroom: an experience. In: Proceedings of the Ed-Media world conference on educational multimedia, hypermedia and telecommunications pp. 1567–1572 (2001)
11. Attewell, J.: From research and development to mobile learning: Tools for education and training providers and their learners. In: Proceedings of mLearn 2005 (accessed on February 10, 2007) http://www.mlearn.org.za/CD/papers/Attewell.pdf
12. Perry, D.: Wireless networking in schools: a decision making leaders. Technology Colleges Trust/ DfES/ Becta (2002) (accessed on February 10 2007) http://www.lea.tameside.sch.uk/schools_grid/ict/wire.pdf
13. Vahey, P., Crawford, V.: Learning With Handhelds: Findings From Classroom Research. SRI International (2003)
14. Gay, G., Rieger, R., Bennington, T.: Using mobile computing to enhance field study. In: Koschmann, T., Hall, R., Miyake, M., CSCL (eds.) CSCL 2: carrying forward the conversation, pp. 507–528. Lawrence Erlbaum Associates Inc. USA (2002)
15. Naismith, L., Lonsdale, P., Vavoula, G., Sharples, M.: Literature Review in Mobile Technologies and Learning. A Report for NESTA Futurelab (2004) (accessed on February 10 2007) http://www.futurelab.org.uk/download/pdfs/research/lit_reviews/futurelab_review_11.pdf
16. Nyiri, K.: Towards a philosophy of mlearning. In: Proceedings of IEEE International Workshop on Wireless and Mobile Technologies in Education (WMTE 2002), Vaxjo, Sweden (2002)
17. Faux, F., McFarlane, A., Nel Roche, N., Facer, K.: Learning with handheld technologies, Futurelab (2006) (accessed on February 10, 2007) http://www.futurelab.org.uk/research/handbooks/05_01.htm

Instructional Technology and Faculty Development: How iWRITE Challenges Course Design and Teaching Methods

Margaret Procter

University of Toronto
15 King's College Circle
Toronto ON M5S 3H7, Canada
procter@chass.utoronto.ca

Abstract. Use of the web-enabled software iWRITE in courses across the disciplines at several Canadian universities demonstrates that implementing instructional technology can lead faculty members to reconsider and reformulate their teaching methods. iWRITE operates in course-specific sites. It supports the integration of writing into disciplinary courses by displaying examples of past student work alongside grading criteria and instructor comments. Instructors unfamiliar with learning theory that validates the use of models may initially fear student copying or distrust the adequacy of student samples as guides, but many have rethought their assignments and strengthened other methods of instruction to take advantage of this method. In co-taught courses, selecting benchmark papers and formulating grading criteria requires examination of disciplinary values and frank discussion of standards. The effort of writing clear comments on the student samples also requires explicit formulation of disciplinary assumptions about the nature of knowledge and the relationship between writing and thinking.

Keywords: instructional technology, writing in the disciplines, faculty development.

1 Introduction and Background: The Design and Aims of iWRITE

Instructional technology can both help and hinder our teaching. We need to know more about how it affects our work—both our larger aims and our practical activities—before we can decide how best to employ it. It may well bring us efficiency in the long term, but only if we know what we want to do with it and how it fits with the principles that shape our work.

As Coordinator of Writing Support at a large research-intensive university, I have had ample opportunity to study the impact of one piece of instructional technology in a range of university disciplines and a variety of teaching situations. My software iWRITE is intended to help instructors demonstrate their expectations for written work, a major challenge in any effort to integrate writing within disciplinary courses.

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 756–765, 2007.
© Springer-Verlag Berlin Heidelberg 2007

It operates in the form of course-specific websites with password-protected access. Each site offers examples of past student work (used with the consent of their authors, and mostly chosen from good or very good papers) that are displayed near a course statement of grading criteria and side by side with detailed instructor comments—thus *showing* as well as *telling* what qualities are valued in written work for the course. iWRITE sites may also include idea-generating Prompter exercises and an online Peer Review component. Except for Peer Review exercises, using iWRITE is always voluntary. Each site has a built-in survey asking for user feedback. The PHP program itself offers instructors a set of interconnected spaces to fill with the samples they select, framed by the criteria statements they formulate, and explained by the detailed comments they write. In other words, instructors have to work hard to create an iWRITE site. I offer them technical help as needed, and also make clear that I or another writing instructor will be available to discuss the content for their sites [1].

In fact, that is one reason iWRITE was designed as empty boxes rather than pre-filled containers—to bring together course instructors and writing instructors to work through questions about the instructional design of specific courses. iWRITE does presuppose two interrelated principles of instruction that may challenge unstated assumptions about power and ownership of knowledge. The basic principle underlying its design is the use of models or examples for learning. iWRITE directly challenges current anxieties about plagiarism and copying by supplying students with relevant texts to imitate in their own work. Logically speaking, the inclusion of those texts in an instructor-sponsored course site makes copying impossible, but fear of plagiarism is still the first reaction of many instructors to the idea of showing student samples. As Howard and Hillocks have both noted [2,3], many faculty members seem to retain Romantic ideas about originality, and may envision students using the samples as templates rather than expressing their own ideas. Others express misgivings about "spoon-feeding" the students by being so directive in the comments and the prompters. I am less sympathetic when they phrase this misgiving as "giving too much away."

A related and sometimes unsettling principle underlying iWRITE is that people can learn from their peers as well as from their teachers. Learning theorists including Vygotsky [4], Lave and Wenger [5], Bandura [6] and Bruffee [7] point out the ways apprentices progress by observing those who are just ahead of them in the process of becoming skilled practitioners. Some instructors, however, feel that giving prominence to students' immature and flawed work threatens disciplinary standards and undermines their own authority. As early theorists of online writing instruction realized [8,9], full use of the new electronic media generates so many different types of texts, including those created by students themselves, that there are literally many authors, not just one spokesperson for the approved type of discourse.

iWRITE was commissioned initially by university administrators who wanted me to "do something with technology" to help teach writing at minimal expense to huge numbers of students. I will not claim even in this panel on "multi-model efficiency" that iWRITE is a magic bullet or a shortcut for the instructional situations it addresses. The efficiency it yields is a matter of longterm investment in course design and resources rather than of immediate payback. I hope to show that iWRITE has been valuable in part because it leads course instructors and writing instructors to think more deeply about principles and practical ways for integrating the learning and

teaching of writing within courses in the disciplines. I see iWRITE as a case study for the difficult but rewarding enterprise that the American Boyer Commission [10] challenged all research universities to undertake.

My analysis is based on my experience across the University of Toronto and my observation of courses in other universities that have adopted iWRITE. This paper will include accounts of resistance, discouragement, and conflict within teaching teams. I will tell stories out of school about iWRITE's role in uncovering flaws in course design and delivery—and sometimes in helping solve them. Many of these challenges and conflicts come from tensions within the instructional enterprise itself, exposing issues that might otherwise have been left unexamined. I will cite real people and real situations rather than merely outlining intentions and hopes, but will use names only where I have sought and been given explicit permission to do so.

2 Reflection and Reconfiguration

The core narrative about iWRITE is brief: setting up course sites makes instructors think about what they are asking students to do, and thus enriches their awareness of course content and teaching methods. One of the early adopters, John Browne, told me exactly that story. He noted that although he had to take six weeks rather than the one or two he had budgeted to set up iWRITE, he gained new awareness by doing so:

> I had to ask myself many questions before I could even begin the task. What really were the goals of the course? How were the goals reflected in the assignments? Were the grading criteria for the assignments consistent with the course goals and with the best practice in the field? How could the Prompters model best practice for novices? What, in fact, *were* the best practices?" (J. Browne, personal communication, January 2006)

Browne noted a specific instance of this rethinking. In setting up a Prompter exercise, he needed to think through a textual feature he had taken for granted, the use of a thesis statement to frame literary essays. As he started creating the Prompter instructions, conferring with me and other colleagues, he found that he wanted to focus on the process as much as the result. His real teaching goal, he realized, was to show that scholarship proceeds by questioning what is already known, developing tentative answers, and proceeding to test those answers by yet further questioning. His Prompter now shows students how to frame focus questions as starting points, then develop them by inquiry and analysis. When Browne looked again at the samples he had chosen, he revised his comments on one unsatisfactory paper to clarify that its problems came from a thesis too neatly stated and too easily supported rather than being examined and tested. The material on his site now reflects a richer and more complex teaching method and a more grounded view of writing as a tool for thinking.

3 Teamwork: Insider and Outsider Contributions

Most iWRITE sites are created for large courses staffed by a lead faculty member and a squadron of teaching assistants (TAs) who mark student work. Ironically, this

arrangement requires the least experienced members of the team to perform a task that has been recognized as among the most challenging in writing instruction [11]. It makes iWRITE valuable as guidance to the changing cohorts of inexperienced TAs as well as to students. Sometimes, in fact, it is TAs or junior faculty who can best define the issues for both audiences.

The first iWRITE sites were team efforts between me and junior instructors in Sociology and Biology who had been delegated to create site content. Each was fully able to identify the good and problematic elements in the sample student texts, but found it hard to explain *why* specific details were suitable or unsuitable, or *why* something was a fine introduction or a good sentence. My role as the writing instructor was to keep asking that question and sometimes to hazard a guess—if only to be corrected—about what exactly was wrong with the word or good about the introduction, why that particular aspect of the piece was worth commenting on, how a student could approach the challenge differently. In the end, some explanations referred to the conventions of genre or discourse, but most touched on the nature of evidence and reasoning in the two disciplines.

I suspect that by the end of that summer the Sociology TA considered my questions something of a nuisance, but he said he was pleased to have learned to be explicit and constructive as a grader. He also told me that he had internalized my type of question in writing his own dissertation. The Biology instructor was already noted for her ability to explain the technicalities of her subject, and she was open to developing the same abilities for explaining writing. As she worked her way through the technical papers chosen for the site, her explanatory comments on flaws of style and reasoning became exquisitely detailed. Facing the need to construct iWRITE comments had brought her to examine her own learning patterns as a student and novice researcher so that she could hear the potential student user asking "What else can I do, then?" Her later work on an iWRITE site for a more advanced course shows the same type of empathy: the comments are now more concise and nuanced, setting the standards higher but still making the advice attainable. It is no surprise to know that this same contributor now runs a private business as an editor of scientific writing.

4 Team Tensions

Issues of collaboration and integration are harder to solve when teaching is done by a team. In theory and sometimes in practice, the discussion and clarification of expectations among team members—whether co-teaching the same section or each teaching different sections—can help improve the focus and consistency of the course. Barbara Rose will outline the process she went through in setting up a major iWRITE site for a multiple-section course, where she managed to obtain consensus about the choice of papers, their grades, the reasons for the grades, and the direction for helpful advice. She may also describe the continuing challenge of maintaining this agreement and of adapting the site to respond to changing views.

In a number of cases, however, I have seen noble aims come to grief because of the inability of instructors to agree on which papers deserve which grades, even before getting to discussions of the reasons why. I have been told of several course meetings

where faculty members taking the lead on an iWRITE project were left feeling shaken at the acrimony of discussion, with colleagues accusing each other of subverting university standards or even undermining the integrity of the discipline because of the way they assessed papers by each other's students. They probably did not thank iWRITE for exposing these disagreements so publicly. But it is worth remembering that students have to move among classes taught by such instructors and face the same discrepancies among their expectations.

One site has taken shape after two years of intense discussion among course instructors in a small department. They initially thought it would be easy to show students what a good essay was in their discipline. In trying to select papers for display, however, they discovered they could not readily explain that point even to each other. They persisted in trying, however, and eventually reached a working agreement on appropriate samples and types of grades and comments. Their site is now available to students in all first-year courses in the department. It contains a rich variety of assignments and a range of comments. Most comments are described as being by "course instructors," but a number are designated as being by "Instructor A" or "Instructor B." Students will see a core consensus, but also become aware that there are different approaches in the department.

Maintaining course sites also requires ongoing consensus and commitment from all the instructors involved. Another story unfolded this year, when one instructor in a multi-section course broke ranks to create a new iWRITE site for her section, working first with a student assistant and then a writing instructor. The pervading issue here seems to be the need for understanding of collaborative roles in such an enterprise. "Whose course is this anyway?" was the recurring question.

The process for the original site had already raised issues about collaborative roles. A writing instructor had realized the potential of iWRITE for her group workshops on exam preparation for a large first-year course where students wrote essay answers worth 30% of the grade in two course exams, but did no writing otherwise. The lead course instructor agreed with the idea, supplied samples, and explained the application of the course grading criteria, but left the writing instructor to write the comments and create a Prompter in which students could draft answers for the workshops. Initial results were encouraging. Students used the site heavily and praised its usefulness, attendance at the workshops increased, and the instructor kept using essay questions. After three years, however, the sample sets on iWRITE were outdated and new ones were not forthcoming.

Enter a new course instructor with a strong commitment to helping students succeed in this high-stakes but low-instruction writing. Julia Richardson wanted to offer her students believable examples of real work done in real conditions, and hoped that a renewed site would encourage them to try a variety of approaches rather than a cookie-cutter formula (Richardson, personal communication, November 2006). She collected over 100 sample pieces, hired a student assistant to help select from them, and enlisted the same writing instructor to mount the material once it was prepared.

Richardson could not change the types of writing tasks in this multi-section course. She could, however, draw on her experience with students and on her knowledge of the qualities that constituted good and poor work in her discipline to overcome the first problem, a false start because the student assistant did not have this experience. He took the collection of pieces and tried to quilt together sections from different

answers to create samples where every part received a clear A, B, or C. Richardson knew that the path to real answers was often bumpier than that. Answers could be excellent in one aspect but poor in others, sometimes from weaknesses in knowledge, but sometimes from misjudgements about timing or proportion. She saw, for instance, that the strikingly good A answer covering five handwritten pages in the exam booklet had cost the author so much time that she was barely able to finish the rest of the exam. It was also useful for students to know that some B answers could have been As if they had included a clear diagram or mentioned an opposing viewpoint, and to see that some A answers still included occasional flaws in language.

Even more of a problem was the tension that arose between Richardson and the writing instructor over how much weight to give to qualities of writing compared to qualities of reasoning. For this site, Richardson drafted the comments herself and asked the writing instructor to look them over and link them to the samples. In pasting text from Word into input boxes, the writing instructor started to feel as if her pedagogical insights were being ignored and that she was being used as a technical assistant. To her mind, some of the answers receiving marks as high as 18 out of 20 were confused and unclear. She was also perturbed that papers presenting clear arguments did not always receive due recognition. For a few days the completion of the site seemed to be in question. Eventually Richardson offered the writing instructor a chance to speak to the class about writing issues, and after a well-received presentation, both realized that their perspectives were not so far apart after all. The comments eventually linked to the sample papers focussed mainly on content issues, but also noted strengths and weaknesses in style and structure that decreased the answers' effectiveness. They pointed out, for instance, that the long A answer not only covered more material than necessary, but was also wordy and overly fancy in phrasing. The excellent argument that received only a B digressed at one point and misunderstood an aspect of the topic. Eventually, both instructors also agreed that the slang and language errors that bothered them were not as important in an exam answer as they would be in a formal paper—but could still be noted in the comments as a concern in academic writing. The writing instructor now has free rein to insert more such comments on papers on the new site. Moreover, her perspective has taken a place among the course topics, and her practice workshops are drawing record crowds of students wanting to hear how developing writing skills can help them demonstrate their knowledge of course material.

5 Making Problems Transparent

One more example will suggest the positive outcomes of such negotiations between writing and disciplinary knowledge—negotiations that become public because of the online distribution of this type of instruction. Even students can have an effect on the design of a resource such as iWRITE, as long as instructors are willing to observe and listen to them. A new first-year Engineering Design course for the cohort of 1100 entering students is a flagship for integration of disciplinary and writing instruction at the University of Toronto—and a huge challenge to teach. This required two-term course has absorbed several courses on Engineering ethics and professional practice, and has replaced a suite of courses that included ESL instruction and technical

writing. Now the course is co-taught by writing instructors appointed to Engineering and other faculty members from a range of Engineering specialties. The course focusses on the process and methods of Engineering projects, including team work, record-keeping, and communicating with members of the public as well as technical experts. Students eventually implement an actual design project with clients from community and non-profit organizations. They thus need to learn the conventions of several types of communication, and they need to manage the processes of iterative document cycling among members of their work teams for peer review as well as grading by various types of instructors. The 50-some instructors who coach students through the various stages of their work also stand to benefit from explicit guidance on what is considered acceptable work in the course.

iWRITE was an obvious choice to support this enterprise. The lead instructor, a research professor in Mechanical Engineering, collected student samples from the initial trial run of the course and enlisted a writing instructor to work with her on appropriate comments. In all, they chose and annotated 35 samples for the 15 different components of the project documentation. As the course unrolled during its first full-scale implementation, the lead instructor created 22 Prompters to take students through the steps of analysis and drafting needed for the various phases of the design project.

As might be expected from such a large enterprise, challenges arose in implementing this course. Students certainly noticed problems and pointed them out—by being restless during lectures, complaining about grades, and making some negative remarks about the course in their opinion surveys at the end of the year. It was also troubling that although nearly 100% of students looked at the samples and comments on iWRITE, only a minority used the Prompters. Mid-year and final surveys given by the course instructors and responses to iWRITE's built-in questionnaire helped define the nature of their concerns, which mainly concerned inconsistencies between Prompter instructions and the grades they had received. Students' abilities to express themselves strongly were quite evident in their angry and often sarcastic responses, well backed up by concrete examples. They needed to be taken into account in the ongoing development of this new course.

Having helped reveal underlying problems with the course, iWRITE has helped address them. Introductory screens, internal comments, and Prompter instructions now all emphasize that the goal of the course is application of Engineering reasoning and the rhetorical principles of audience analysis; they are not simply templates for written products. The Prompters have also been recast to make more explicit their applicability at different stages of the course, reflecting the increasingly complex material and increasingly professional roles students have to play in the second-term design-project iterations. And they have been thoroughly checked for consistency.

I found myself volunteering to take on this work of revision for the second year of the course, and I thus ended up reading through course material, cornering instructors at social occasions, and dropping into lectures. I followed up by sending revised Prompters to these contacts for their information and comment. In all this work, I found I had much to learn myself about collaborative writing and iterative revision. I have now accepted the use of "shall" where I would have insisted on "will" or even "would" (because "shall" expresses a legal constraint or promise), and have seen the need to eliminate such standard academic hedges as "usually" and "perhaps" (because

they are imprecise). I have now also accommodated my advice on revision to the Engineering practice of retaining pre-set section headings and then adding and changing material and style within them rather than reorganizing freely.

I am still in the middle of this process, and can report that it is stimulating, frustrating, and productive in about equal measures. Students are using iWRITE as heavily as last year and, as they move into individual work on the second of their three project iterations, are also starting to turn to the Prompters for guidance.

6 Resistance and Adaptation

Although iWRITE is a success in terms of numbers using it (over 10,000 at this 72,000-student university) and the enthusiasm of students who have access to it, I still meet much faculty resistance to adopting it for their own courses. The fear of student copying, or of undue imitation, remains a strong deterrent to many who hear about iWRITE from their colleagues or at my faculty workshops. To deflect it, I like to volunteer that I know of only two cases where students got into trouble by misusing iWRITE material. In one very large first-year course, a student handed in as his own work the exact same text as one of the iWRITE samples and was immediately charged with plagiarism. His excuse was that he couldn't do the assigned work, so he chose a D paper because he knew that was all he was worth. In the other case, a student copied an iWRITE sample into her own word-processing file and then changed nearly every word to fit her topic. But she forgot to change the title, and the TA grader then noted how close her work was to the structure and reasoning of the sample. This student was given another chance to write the piece for herself, and was urged to use the Prompters to help formulate her own thinking on the topic.

It is the online presentation of the sample material that makes those kinds of copying possible, but the internal conditions in the course—both the material conditions of huge classes and lack of student contact, and the nature of assignment design—that make them attractive to at least a few students. If instructors have been inclined to recycle a few standard assignment topics, that practice will be made impossible by iWRITE. My most unsettling experience with iWRITE came when a large first-year science course developed an elaborate site to show a sequence of increasingly complex set of lab reports, but then was unable to change the lab topics by the next year, as intended. Students found it hard to come up with different words for their lab reports than the ones on the site, and TAs were faced with the additional burden of trying to assess adequate degrees of paraphrasing. This site was withdrawn from student use after the first year, but can be made available again once the lab topics are changed. Meanwhile, the current group of TAs has requested access so they can see examples of "benchmark" papers and of good commenting practices. They are much happier this year, though their students are still not fully served.

Other practical difficulties remain with the TA role in courses that use iWRITE. In both the cases just mentioned, the TAs in the course knew the iWRITE samples well enough to catch the problems immediately. There may well be instances where graders are not familiar with the iWRITE material, perhaps after a few years of use when the resource has started to be taken for granted, and when old assignment topics may have crept in again. The same situation could arise with new instructors in

multiple-section courses—few of whom will have the energy of Julia Richardson to create their own new sites. TAs and new course instructors also need to read through and master the material on the course site. Here the increase in preparation time is recouped by a decrease in student complaints and by the usefulness of iWRITE as a stock of "teachable moments" for use in advising both TAs and students.

Evidence is also available to defuse instructors' misgivings that students may limit their efforts to imitating approved structures or methods of reasoning. Monitoring of student work in several courses has shown increases in synthesizing of information and more attempts at critical thinking—imitation of higher-level elements rather than duplication of superficial aspects. Students' comments about their experiences of iWRITE [12] show an acute awareness of their liminal situation: they appreciate the chance to look into the world of the successful students who wrote the A and B papers on the site, and then into the dark teacher's room where the instructors write their comments. The most common response to our online questionnaires is that students no longer feel "clueless," as one of them put it. They realize that there are ways of solving the problems posed by course assignments, and that success is within reach. The next most common response is that they realize they should aim higher than they had thought. As one student told us, "I see that I can write a C paper now and might be able to write a B paper, so I should aim at an A paper."

In spite of the satisfaction of hearing such remarks about positive inspiration, I have also had to recognize that samples are equally valuable when they warn against problems and are used for direct instruction. Indeed, iWRITE always positions instructors as the assigners of grades, and it uses their voices as the creators of those guiding comments and the Prompter instructions. It is only fair, after all, to help students avoid injuring themselves on the machinery of their disciplinary apprenticeship. Pointing out flaws in structure and reasoning even in good papers is one way to reassure students that perfection is not expected—especially when the flawed papers and the negative analysis has no personal bearing on the student using the iWRITE site. A number of sites have created special sections to show rewritten sentences or passages as a way of teaching style and coherence in their disciplines. Others comment on how a B paper could have been an A with specific types of revision. In courses where conventions of disciplinary writing are entirely new to students, comments on a C paper can demonstrate the effect of sincere misunderstandings of the task. Comments on surface details of format, including referencing systems, have their own use too—in fact are one way to avoid harping on such matters in class or in repeated marginal notes on students' actual texts. The comments on the iWRITE samples are also useful as models to TAs and writing-centre instructors who are themselves novices in knowing and explaining why some strategies are not valued in particular disciplines.

One final set of evidence leaves me encouraged that even those who do not use iWRITE themselves have gained something from knowing about it. For every course that adopts iWRITE and pulls together the resources to set up its own site, I hear of at least two others that begin to show and discuss student samples in class. The use of grading criteria is already common because of rules for accountability, but I also hear from many people who have expressed interest but not adopted iWRITE that they are starting to make those criteria statements a focus of class discussion and practice exercises. The next stage is often a revision of the statements to make them clearer

and more precise. And a few instructors have come back to say that these teaching strategies were so useful that they were now ready after all to set up an iWRITE site.

7 Conclusion

I sometimes have the heretical thought that iWRITE could be replaced by the Insert Comment function in Word or the Commenting Tool in Adobe Acrobat, backed up by excellent classroom instruction and generous office hours—if instructors only had the time and inclination to interact personally with all their students. But that "if" condition cannot be taken for granted at my university or most others. Thus the advantages to students of all-hours access to the resource make the work of constructing a site worthwhile in practical terms. Nearly all instructors who have created iWRITE sites for their courses have continued to update their sites and use them year after year. The hard work of reconceptualizing approaches and rethinking specific strategies has gained them considerable payback in terms of pedagogical efficiency and effectiveness, and I believe has also made the effort worthwhile in terms of their self-knowledge and satisfaction with their work. Sharing in that hard work has certainly enriched my teaching experience and the ways that I have been able to help integrate writing into disciplinary courses.

References

1. Procter, M.: iWRITE Web-enabled Software (Interactive Writing Tutorial Environment) Notes for Course Instructors (2007) http://www.utoronto.ca/writing/iwrite/demo.html
2. Howard, R.M.: Standing in the Shadow of Giants: Plagiarists, Authors, Collaborators. Ablex, Stamford, CT (1999)
3. Hillocks, G.: Ways of Thinking, Ways of Teaching. Columbia Teachers College Press, New York (1999)
4. Vygotsky, L.S.: Mind in Society. Harvard University Press, Cambridge, MA (1978)
5. Lave, J., Wenger, E.: Situated Learning: Legitimate Peripheral Participation. Cambridge University Press, Cambridge (1991)
6. Bandura, A.J.: Self-Efficacy: The Exercise of Control. Freeman, New York (1997)
7. Bruffee, K.A.: Collaborative Learning: Higher Education, Interdependence, and the Authority of Knowledge, 2nd edn. Johns Hopkins University Press, Baltimore (1999)
8. Warschauer, M.: Electronic Literacies: Language, Culture, and Power in Online Education. Erlbaum, Mahwah NJ (1999)
9. Hawisher, G.E., Selfe, C.L.: Global Literacies and the World-Wide Web. Routledge, New York (2000)
10. Boyer Commission on Undergraduate Education: Report on Educating Undergraduates in the Research University: A Blueprint for America's Research Universities (1998) http://naples.cc.sunysb.edu/Pres/boyer.nsf/
11. Ferris, D.R.: Response to Student Writing: Implications for Second Language Students. Erlbaum, Mahwah NJ (2003)
12. Student Comments on iWRITE from Online Questionnaires 2005-6 and 2004-5 (2007) http://www.utoronto.ca/writing/iwrite/comments05-06.html

Technology-Mediated Provision of Models: Observational Learning for First Year Students

J. Barbara Rose

Woodsworth College, University of Toronto
119 St. George Street
Toronto ON M5S 1A9, Canada
jbarbara.rose@utoronto.ca

Abstract. First year students coming into university are often puzzled by the expectations of university discourse. For students in the Academic Bridging Program (ABP) directed at mature and under-prepared students this is especially manifest; these same questions are often expressed by students in traditional first year university classes. This challenge has been assisted by the use of a web-enabled software program, iWRITE which displays examples of past student work of essay assignments along with detailed comments by course instructors. The subsequent success of students in acquiring not only structures and vocabulary to help them navigate this new world of academic discourse but also a heightened sense of self-efficacy supports the concept of observational learning by the technology-mediated provision of models.

Keywords: instructional technology, observational learning, self-efficacy.

1 Introduction and Background

In my capacity as Lecturer at Woodsworth College at the University of Toronto, I teach ENG185Y, Introduction to the Study of Literature, in the Academic Bridging Program; in addition, I teach a First-Year Seminar Course, HUM199Y, Fatal Attraction: The Lure of the Villain in Literature, as a regular Arts and Science credit course. The suite of 199Y courses across the range of disciplines in that faculty is meant to introduce high-school graduates to university-level discourse and to give them a rare small-class experience. I speak from a double perspective as not only am I a faculty member teaching literature courses, I also have many years of experience as an instructor within various writing centres across the university and thus have had an opportunity to work with and observe students from many disciplines in their attempts to address and complete writing assignments. One of most frequent comments I have heard from students over the years is the plea to see a copy of "what an essay looks like". My students are very anxious about writing and one of their first priorities is the desire to improve their writing skills as quickly as possible. This attitude toward writing is reinforced by a recent Harvard study where "of the skills students say they want to strengthen, writing is mentioned three times more than any other" [1]. It is this aspiration and my subsequent exposure to Margaret Procter's technologically mediated program of essay examples of past student work alongside

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 766–775, 2007.
© Springer-Verlag Berlin Heidelberg 2007

detailed comments by course instructors which motivated me to first implement iWRITE for my Academic Bridging Students at Woodsworth College. I will not duplicate what Margaret Procter has previously outlined about the iWRITE program in her presentation; more specific information is available on http://www.utoronto.ca/writing/demo.html.

Woodsworth College is a relatively new college established as part of the University of Toronto in 1974 to address the needs of part-time students in undergraduate programs. It has also provided opportunities for non-traditional students to pursue degree studies. The Academic Bridging Program (ABP) enables mature students (in this definition, "mature students" must be at least 20 years of age without any prior university courses) to succeed at university. It offers courses designed for students who have been away from formal education for some time (or who perhaps did poorly at high school) and who need to bridge the gap between their prior education and the requirement of first year university courses. The subjects offered are Introduction to Literature, Contemporary Canada, and Canadian History—all classic Humanities courses requiring extensive reading and much analytic writing. Students who successfully complete an ABP course gain a full credit towards their degree studies in the Faculty of Arts and Sciences at the University of Toronto. Recently, in 2001, Woodsworth has also started to take full-time direct-entry students from high school, and to offer courses tailored for the different type of transition they need to make from secondary to university studies. Woodsworth is currently one of the largest colleges at the University of Toronto, with approximately 6,700 students registered, both full-time and part-time, including international and visiting students.

I have used iWRITE for the non-traditional students in the ENG185Y Academic Bridging Course for three years, and to date have garnered encouraging and remarkable results, which I will discuss in this paper. However, I have postponed setting up a full version for the high-school graduates in the HUM199Y course in order to assess the students' needs before launching the site. Although I had some samples available from two years of teaching this course, I wanted to gather more meaningful examples and give the students (mostly science students) a broader range to look at. Whereas I assign standard tasks of analysis and interpretation in my ABP course in order to induct students into the basic patterns of university discourse, in my HUM199 course I have developed assignments asking students to bring knowledge from their other courses and apply it to the interpretation of literature. Some of the students have already been able to absorb this challenge and have developed exciting and creative expressions, bringing their course work from History, Cinema, Sociology, and Psychology into their literature essays. I look forward to showing five or six of the best of these pieces to future classes of HUM199 students to help them see the range of what is possible in academic writing. I have already started with a somewhat different site entitled "Don't Make These Mistakes" highlighting examples (taken from previous years' student essays) of some of the specific components such as overwritten and bloated introductions, improper use of quotations, grammar and sentences problems, and sloppy conclusions. My aim is to focus on frequent student errors in order to deter my students from resorting to this type of usage in their short essays of which I will see a draft form and they then will re-submit two weeks later. My intention is to introduce students to the full site with examples of longer essays after the return of their essays from their first short assignment in order to gauge and

to monitor effectiveness of the site. The experience and purpose of the HUM199 iWRITE site differs substantially from the ENG185Y site.

The tailoring of iWRITE to the specific needs of students in ENG185Y, the Academic Bridging course that I teach, is the key to its success as an instructional resource. The types of students applying to the Academic Bridging Program are unique. For the older among them, coming to university is the fulfillment of a life-long dream. For the younger ones, it is a second chance at a university education: many may have graduated from high school with poor marks, or dropped out before graduating, and have been working in low-paying dead-end jobs.

These students bring into the classroom many levels of skills and preparation. They also bring a great deal of prior knowledge and experience. Although these skills and knowledge are not necessarily academic, it is the instructor's challenge to harness them. Many students are intimidated by the buildings, the classroom, the lecturers, their classmates, the course content, and above all, by the writing requirements. Instructors in these Academic Bridging courses not only have to teach course content, they also need to teach reading, writing, and thinking skills as well as study and time management skills. In addition, we need to instill a level of self-confidence that will allow them to tackle the assignments with success. It is this level of self-confidence which leads to a sense of perceived self-efficacy about their abilities to complete the assigned writing tasks at a university level; this self-efficacy pays an important role in their learning process [2]. We also need to help them become part active members of a new "community of learning," fulfilling Vygotsky's theoretical dictum that the "higher mental functions are socially formed and culturally transmitted" [3]. Thus we must recognize that social interaction plays a fundamental role in the development of cognition. These students soon learn that the range of skills that can be developed is far greater with guidance, modeling, and peer collaboration than if they tried to do the assignments by themselves. Our results bear out this out.

Courses in the ABP are heavily loaded towards written assignments, writing in class, and a final three-hour exam. Students are usually required to write at least 8,000 words during a course, usually consisting of four in-class assignments and four essays, one of which must be a major research essay. This can be a daunting requirement for many of the students, some of whom have not written an essay in 30 years, or in some cases ever, though they may have written in business situations. One student confessed that the only writing she does is a weekly shopping list. She felt the task of writing the first essay (a 750-word piece) was overwhelming.

Classes for these non-traditional students are kept fairly small compared to those in the regular program. They are usually capped at 50, although the university is attempting to improve the first year experience for students and this year these courses have been capped at 40. Thus instructors can provide these students with far greater individual attention than is possible in the normal 100+ class sizes in undergraduate programs. These students also have many other scaffolds to support them: the services of the Academic Writing Centre at Woodsworth College, a set of workshops on academic skills tailored for their likely questions and concerns, and full access to the Woodsworth Computer Lab to work on their assignments. Since many students are older than the usual undergraduate (some in their 70s and 80s, for example), they are not always comfortable with computers and do not always have easy access to them at home.

The students in my ABP course come together in September with very little in common save the fact they are in my classroom. They are very different from the 24 students in my HUM199Y course, who are homogeneous in age and in high-school experiences and skills levels. For the ABP students, the classroom meetings and the optional out-of-class support activities are intended to create what Lave and Wenger call a "community of practice" in which they are "apprentices" [4]. There they share common activities such as reading specific texts, writing specific assignments, discussing certain plays, poems, works of fiction, etc. These common activities give the ABP students a sense of joint enterprise, and more importantly, a sense of identity and belonging.

And yet the academic culture is still not comfortable for many. Students have to enter writing situations where genres are already set and where they need to fit into certain specific expectations if they are to succeed. Students must learn the language and culture of this new community, and they need to find the relationship between themselves as participants, the activity, and the environment. Current theorists stress that all human activity—including writing and learning—is situated in specific social situations and requires that people act from a role or position. In this case, the role is that of a student, an unaccustomed or even a threatening one. As Bruckner articulates, everyone inhabits multiple roles throughout the day and throughout life [5]. For the high-school graduates in my 199Y course, these identities are often assumed without question or problem. For ABP students, there are two sets of challenges: to gain human capital in the form of education that will improve their job credentials and employability, and to gain social capital in the form of emotional support and encouragement from their fellow students. Instructors thus need to capture both types of motivation, both the individual ambition and the sense of group solidarity and shared enterprise. Bruckner describes this process with her City Writers group when she says that "we might see our task as one of providing newcomers with the resources they need—materials, time, a sense of the overall goals of practice, history, access to expert peers, motivation, protection—to become full participants in the community of practice" [5]. This is certainly my aim in all my classes, but especially in the Academic Bridging Program.

iWRITE is one of the best resources available to students in my Introduction to Literature classes. It helps them fulfill exactly the move that Lave and Wenger define as crucial, moving from "legitimate peripheral participation" [4] closer to the centre of their new academic world by achieving successes in performing the key apprentice activity of writing, that is, using the new language and way of thinking of the academic community. Moreover, it shows them the successes of other apprentices, former Academic Bridging students, as encouragement. As Pajares points out, Bandura (1986) emphasized that comparisons with peers are important determinants of self-efficacy beliefs [6], and no matter how frequently I may rely on studies and statistics, students are skeptical, yet when I provide actual student sample essays on iWRITE, they understand much more quickly what I am attempting to show them. One of the most frequent student requests to instructors is a plea for more detailed information about expectations for the written assignments. For students, the sample essays represent a point of reference or a standard by which students can measure themselves; without these standards, students are left in what Bandura labels “foggy ambiguity” [7]. In addition, this ability or “personal efficacy in mastering tasks” acts

as a great motivator and encourages further effort [7]. The explanations on iWRITE give these students not only examples of students' essays, but also instructors' comments clearly outlining what is appropriate and what falls outside of standard academic discourse.

Over the years, as an instructor in the Woodsworth Academic Writing Centre, I have collected student essays from many disciplines so that I can provide students with examples when they ask to see an "essay." iWRITE now helps me with that supportive activity for my course, enabling students to visit on-line as often as they like to view examples of various types of essay assignments done for previous sections of the course, as well as samples of in-class writing assignments and short-answer questions. For my students, iWRITE supports Bandura's concept of "observational learning." I like his explanation:

> Learning would be exceedingly laborious, not to mention hazardous, if people had to rely solely on the effects of their own actions to inform them what to do. Fortunately, most human behavior is learned observationally through modeling: from observing others one forms an idea of how new behaviors are performed, and on later occasions this coded information serves as a guide for action. [8]

This "observational learning" or "modeling" is what iWRITE does for my students, allowing them to view and then try out structures, vocabulary, and models to help them understand their new world of academic discourse. As students complete their essay assignments and are successful, they realize that "they are capable of performing well" [2] and this, in turn, enhances self-efficacy.

Another advantage for students is the time factor. Bandura comments that this method of teaching allows students to learn "integrated patterns of behavior without having to form them gradually by tedious trial and error". The method of instruction is important for "both development and survival" [8]. These ABP students need this course for entrance into regular undergraduate classes and their motivation to work hard is usually much higher than traditional students in first year courses. Because of their many domestic commitments (family, employment, etc.), they take advantage of any resources that will save them time in acquiring the skills they need since they do not have the luxury of "tedious trial and error."

2 Initial Resistance to the "Modeling" Function of iWRITE

Since the ABP has many sections and many instructors, I was acutely aware that I would need the co-operation and support of the other eight or so instructors of the Introduction to Literature course to launch this resource. This has not been easy to accomplish. Some of the instructors (many of them also working elsewhere and not regular university faculty) were not overly enthusiastic about the use of computer-assisted learning and were resistant to having their students use the program or even contribute essays to the database. Their underlying fear seemed to be that showing students good examples would "give away" the skills that they wanted them to develop by trial and error.

A few students also resist using iWRITE because of their romantic idea of academic writing. Some always start the course saying that they do not want to be

"influenced" by the opinions of others; they want their ideas to be "fresh" and "original." Some of these students express disdain at the idea that they might be asked to model their writing on that of other students. A few always refuse to log on to iWRITE, even after I give notable examples of the advantages of observing and using samples of others' work: Beethoven adopting the classical forms of Haydn and Mozart, for instance, and Shakespeare using and adapting his sources.

3 Advantage of iWRITE: For Students

Once they try iWRITE, however, nearly all students see the obvious benefits from using the resource to support their learning in the program. Students' resistance to using iWRITE is weakening, however, as students hear from each other that iWRITE really does help. Last year over 60% of the students in this high-attrition course used iWRITE; in my class section nearly everyone did. Perhaps the chief advantage is that students can view the essays whenever they want and as frequently as they like, coming back to make comparisons or look at sections of specific papers more closely. Bandura comments that people usually refine modeling through "self-constructive adjustments on the basis of informative feedback from performance and from focussed demonstration of segments that have been only partially learned" [8]. My students bear out this theory, noting that they like to work through iWRITE at their own speed and skills levels; I note that these resisters tend to turn to iWRITE after getting back an assignment where they did not do well. Once they try out the strategies they have seen demonstrated in iWRITE, they are also encouraged to study its examples and advice further—a self-feeding cycle of support and further learning.

4 Advantages of IWRITE: Specific Student Experience

One of my mature student who attended a focus group organized by Margaret Procter to gather feedback on iWRITE after its first year of use exemplifies the nature of this cycle. She arrived at the focus group meeting waving two exam books, eager to tell all of us about the effects of iWRITE on her work. It was solely responsible, she said, for raising her grade by 15% between the first in-class assignment and the second—the key to her having stayed in the course and eventually succeeded in gaining university entrance after a hiatus of at least 10 years and some very difficult life experiences. The papers she brought with her do demonstrate that iWRITE played a large role in that leap of confidence and skill, though we will not claim that it was the only factor. However, as Pajares notes "confident students are also likely to feel less apprehensive and have stronger feelings of self-worth about their writing skills" [6]; the student's subsequent progress in the course reflects her increasing feelings of self-confidence.

The student had worked hard to prepare for the first in-class writing assignment, searching out criticism on *Oedipus Rex* and *Othello* and attempting to appreciate their stature as masterpieces. However, her answer to "identify and comment" question on passages from those works garnered her only a 62% grade, a bare C- that would not help her gain university admission.

In writing the next in-class assignment a month later, this student wrote a very competent analysis of a sight poem, describing its mood accurately and setting out the techniques used by the poet to create it; I assigned a mark of 77% for her attempt. According to the student, it was studying the iWRITE examples of successful in-class essays, including explications of poems, that gave her the confidence to try literary analysis on her own. The self-directed activity, and the identification with other successful students, made all the difference to her sense of where she belonged in the course. It seems evident that the student's confidence, as well as competence, increased when she was given what Pajares calls "process goals" [6] such as specific strategies via instructor feedback (focus, analyze text) and the use of iWRITE that the student can use to improve her writing.

5 Advantages of iWRITE: General Findings

Another interesting student comment highlights a facet of iWRITE that had not occurred to me. In the April 11, 2006 exit survey (Appendix A) on iWRITE usage, a student answered the question: "Would you recommend the use of iWRITE for future Academic Bridging students?" and "Why?" in the following way: "Without a doubt. It is a useful guide, particularly for prohibitively shy students. It is also a secure, professor-approved site." Some students feel intimidated speaking in class and using iWRITE provides them with ready access to information but with protective anonymity.

A more banal statistic also supports the idea that using iWRITE helps students succeed in the Academic Bridging Program. The ENG185Y site was created for the Introduction to Literature course three years ago and has been used heavily since then. iWRITE sites have not yet been created for the other two Academic Bridging courses on historical topics—in spite of student pleas, including their tendency to show up at demonstration sessions for the Literature students in the hope of catching a glimpse of real papers, even if they are not for the right course. The contrast provides some numerical evidence that iWRITE helps provide an advantage to its users. From the original Introduction to Literature classes of 50, about 35 students now turn up to write the final exam. The courses in the other two subjects also start at about 50, but only 20 or so students take the final exam. Without the confidence to think they might succeed, the other students have given up their chance to enter the learning community. I am currently in the process of meeting with the instructors from the other Academic Bridging courses, Contemporary Canada and Canadian History, setting up guidelines and sorting through examples of essays that will prove suitable for students' access and use.

6 Advantages of iWRITE: For Instructors

For instructors, once they overcome their initial resistance, iWRITE also serves many purposes. It helps especially in preparing students to write the early papers in our courses, when they need to learn some of the mechanical aspects of essay-writing (such as page format, referencing methods, and paragraph structure) as well as the

vocabulary and content of academic discourse. I am certainly glad myself not to have to mark the same errors of format, style, tone and logic as often as I used to. Bandura notes that "it is often difficult to convey through words the same amount of information contained in pictorial or live demonstrations" [7]. His dictum is borne out by student successes in the basic elements of format, which we as academics take for granted, but those unfamiliar with school settings do not—even when they are specified as part of essay instructions. iWRITE shows students such basic concepts as double spacing, paragraph indents, length of paragraphs, referencing format, etc., with enough comments to show that instructors really do take all that seriously. More importantly, students can see what we mean by having an explicit thesis, a pattern of argument, topic sentences, and paragraph transitions. The large number of students' comments on this very aspect attest to this function of iWRITE (see Appendix B). I could show you a typical example: a page covered with short choppy paragraphs, the opening of an essay attempting to say everything possible about *Oedipus Rex*, and succeeding only in re-telling the story incoherently. After exposure to iWRITE, the student began her second version with two substantial paragraphs framed by topic sentences. Her introduction now set out a nascent argument supported by specific references to the play. The writer had passed the first hurdle to being taken seriously by others in her academic community. Now she could proceed with confidence to develop more advanced capacities for academic thinking and communication.

7 Conclusion

iWRITE with its technology-mediated provision of models has proven to be a valuable resource for my ABP ENG185Y students, since it provides them with examples of success by their peers in the tasks set by their own course. These students gain confidence and there seems to be a strong link between self-efficacy beliefs and achievement outcomes [6]. The use of iWRITE for my HUM199Y Fatal Attraction course will match my assignment design there, providing students with examples showing how scientific and other types of academic knowledge can be integrated with the interpretation of literature. As Bank emphasizes, we can help students understand the writing process by providing them with examples of good writing [9]. I look forward to helping instructors in the other two ABP courses set up sites for their courses too, responding to student demand for the same type of resource as their peers in the Introduction to Literature course.

Ultimately the ABP students gain proficiency and self-efficacy in this "community of practice"—as well as a credit for passing the course—through the process of identifying with the valued practices of the university. We know from many studies that students "with positive expectations that result from a strong sense of confidence approach tasks with optimism and continue to strive" [6]. Armed with their "apprenticeship papers," these Academic Bridging students go onto other courses, once again learning what is required in their new courses, but with a better idea of the writing expectations and practices that will guide them on their journey at the University of Toronto.

8 Future Research

A direction of my future research during my sabbatical in 2007 is to follow students as they navigate their way through university programs in different disciplines to analyze changes in writing over time to assess the impact and significance of this initial iWRITE experience.

References

1. Light, R.J.: Making the Most of College: Students Speak their Minds. Harvard University Press, Cambridge, MA (2001)
2. Schunk, D.H.: Self-efficacy for Reading and Writing: Influence of Modeling, Goal Setting, and Self-evaluation. Reading and Writing Quarterly 19, 159–172 (2003)
3. Vygotsky, L.S.: Mind in Society. In: Cole, M., John-Steiner, V., Scribner, S., Souberman, E. (eds.) Harvard University Press, Cambridge, MA (1978)
4. Lave, J., Wenger, E.: Situated learning: Legitimate Peripheral Participation. Cambridge University Press, Cambridge UK (1991)
5. Bruckner, T.: It Just Sort of Evolved: Negotiating Group Identity Among Writers. The Writing Instructor (2004) Retrieved June 17, 2005 from http://www.writinginstructor.com/essays/bruckner-all.html
6. Pajares, F.: Self-efficacy Beliefs, Motivation, and Achievement in Writing: A Review of the Literature. Reading & Writing Quarterly 19, 139–158 (2003)
7. Bandura, A.J.: Self-efficacy: The Exercise of Control. W. H. Freeman and Company, New York (1997)
8. Bandura, A.J.: Social Learning Theory. Prentice-Hall, Englewood Cliffs, NJ (1977)
9. Bank, C.: Reading and Writing Taught in a Sophomore Course in Plate Tectonics. Journal of Geoscience Education 54(1), 25–30 (2006)

Appendix A

ENG185Y
iWRITE USAGE
Survey -- 2006

1. Did you use iWRITE? No _____ Yes _____
2. If yes, how often?
3. If not, why not?
4. In what ways did iWRITE help you with your writing skills?
5. What changes would you make to iWRITE?
6. Would you recommend the use of iWRITE to future Academic Bridging Students? Why?
7. In your opinion, is iWRITE difficult to use? Why?
8. What resources would you recommend for students coming into the Academic Bridging Program.

Appendix B

ENG185Y
iWRITE USAGE SURVEY
April, 2006

Question #4: In what ways did iWRITE help you with your writing skills?

- Acted as a template.
- Served as examples for style, content, and mechanics.
- iWRITE helped me see other essays and so I could base my essays on these. I could see exactly what the prof. was looking for. Actually, it helped with an in class essay, and I received a good mark!
- iWRITE shows me what is expected of me for each assignment. It also helps me see different ways I can structure my essay.
- iWRITE showed me structure, knowledge on works cited and vocabulary; it is better than other sites.
- With the poetry essay it showed me a well-defined structure for discussing the two poems. With the research essay it cleared up minor questions regarding MLA format.
- Provided me with examples of appropriate essays, grammar, structure etc. Also provided examples of BAD essays, grammar, structure, etc.
- It gave me ideas of how to approach topics and what an essay of a certain mark looks like. I also looked at names others used as secondary sources.
- The essays posted on iWRITE served as concrete examples that guided me significantly.
- Understanding the assignment questions, and help on thesis and structure of essays.
- I was able to model my essays based on the ones posted and understand what prof. wanted. Gave me confidence that I understood all implications of the research essay.
- Helped me not to panic.
- It helped to demonstrate format for types of writing I was not yet familiar with.
- iWRITE allows you to parallel your paper; it is a tool to use as a guided format. It is not necessarily something that you copy, but you use the "model" for the format and content that you can expand in your own paper.

English for Academic Purposes in Israel: Perceptions of E-Learning from the Perspectives of Learners and Teachers

Bella Rubin and Helen Sarid

Tel Aviv University, Division of Foreign Languages, 8 Sasha Argov Street
Tel Aviv 69620, Israel
Tel Aviv University, Division of Foreign Languages, 4 Groneman Street, Tel Aviv
brubin@post.tau.ac.il, hsarid@post.tau.ac.il

Abstract. This exploratory study investigates, from the perspectives of learners and teachers, how e-learning is implemented in the domain of English for Academic Purposes (EAP), particularly in the teaching of reading and writing skills in universities and colleges in Israel. The study adapts Ben Shneiderman's [1] view that learner-centered measures should be used to evaluate educational progress. The study was conducted in a natural setting, using a grounded approach and qualitative interpretation of data. Teachers who used e-learning responded to open-ended questions about their practices. Their students responded to open-ended questions which explored their perceptions of learning outcomes in academic language courses. To elicit in-depth reflection, volunteer students and teachers were interviewed. The findings may lay the groundwork for future studies exploring how e-learning affects students' learning strategies, teacher/student relationships, the very nature of the classroom, to name but a few theoretical and pedagogical issues encountered.

Keywords: EAP, perceived learning outcomes, e-learning, blended learning.

1 Introduction

"The sage on the stage is becoming the guide on the side, and collaboration is replacing competition." [2 p.vi]

In recent years institutions of higher education (HE) in Israel have been using learning technologies in delivering courses, not only in the sciences but also in the humanities and social sciences. Fast changing technology enabling construction of intra/interuniversity portals has enhanced widespread use of web-supported instruction and e-learning, but this transformation has also raised issues related to the quality, efficiency and flexible nature of such instruction. A study of how one major university in Israel has implemented blended learning across disciplines [3] has provided insights into the process of moving toward campus-wide web-supported courses. The researchers found that lecturers adopted the technology readily but that the pedagogical range of implemented activities was still limited. The current work focuses on how e-learning is being implemented in the domain of English for Academic Purposes (EAP), as taught at several universities and colleges in Israel.

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 776–785, 2007.
© Springer-Verlag Berlin Heidelberg 2007

Ben Shneiderman's informative essay, "The New Education," [1] provides teachers with a philosophy for how to take full advantage of what he calls the new technology. According to Shneiderman, e-learning that takes into account the four stages of human activity, collect-relate-create-donate, facilitates active learning and inquiry-based learning, two of the important principles upheld by many educators. In applying these ideas to language learning through technology, Debski [4 p.39] advocates "…learning environments in which students can easily exercise their creativity, collaborate on projects of interest, engage in goal-driven activity, and combine learning a language with reflection about language and learning strategies."

With these views in mind, we explored the experiences of learners and teachers engaged in e-learning in a variety of EAP programs focusing on reading and writing skills. Undergraduate students in Israel's six universities and numerous colleges must attain advanced level English reading skills (and in some institutions, writing skills) since much of the reading material is in English, though instruction is in Hebrew. Some students are exempt from EAP courses based on their English scores on university entrance examinations; others have a year to complete required EAP courses which bring them to exemption level. English is taught as a required foreign language in primary school from grade three. Thus, English plays a major role in HE and also in Israeli society and the business community. English is a second language for native Hebrew speakers and a third for native Arabic and Russian speakers.

In this study we do not aim to evaluate learning outcomes or the efficiency of teaching methods or practices but rather to describe the paths that have been taken by different institutions and individual teachers and learners in their engagement with e-learning environments and to relate some of the reflections of these participants.

How is E-learning Defined in this Study? What are its Aims? For this study, we included a variety of computer-supported activities which we defined as e-learning:

- using the Internet in class or at home
- sending email communications to teachers and among students
- working on computer-based tasks related to course content in a lab or learning center
- accessing course materials, announcements, homework assignments in web-supported courses through a university portal

According to Nachmias and his coworkers [3 p.375], in their description of the Virtual TAU project at Tel Aviv University in Israel, "The project aimed to initiate and stimulate a process by which more faculty members will gradually use the Internet to enrich prevailing learning processes and make instruction more efficient and flexible." We would like to adapt these aims to EAP course goals and include any e-learning activity (such as listed above) used to enhance the learning of reading and writing skills in an academic program.

In addition, though the current study does not aim to evaluate e-learning practices or outcomes, for data analysis, we found it helpful to keep in mind the four benefit components defined by Cohen and Nachmias [5 pp.84-86] in their quantitative model of cost effectiveness for web-supported academic instruction. The benefit components are: "improving instruction quality, improving affective aspects, increasing efficiency of teaching and learning processes, and facilitating knowledge management."

1.1 The Study

This study was conducted in a natural setting using students and teachers participating in English courses required for exemption level in various undergraduate degree programs. A grounded approach to data collection was used. The data was collected through self-reporting questionnaires and semi-structured interviews and analyzed qualitatively [6]. This approach enabled us to explore the perceptions of teachers and learners engaged in e-learning activities which are integrated into their EAP courses.

1.2 The Research Questions

From the perspectives of teachers and learners, how is e-learning being implemented in EAP courses in universities and colleges in Israel?

- What progress in reading and writing in English do students perceive they have made as a result of their use of e-learning in EAP courses?
- What other outcomes do students perceive as a result of their e-learning experiences in EAP courses?
- What are teachers' perceptions about e-learning in their language courses for academic purposes?

2 Method

The Sample. There were 100 undergraduate students from various disciplines (e.g., biology, physics, computer science, Middle East studies, education, art) in the study, all participants in intermediate and advanced level EAP courses taught by six teachers at a major university in Israel. In addition, three teachers from other universities and colleges participated by either responding to questionnaires or undergoing interviews. One of these teachers taught EAP (advanced level), one taught French for Academic Purposes, and one taught business English in an academic program and teacher training for Spanish teachers (advanced level). All courses included various e-learning components, as listed above, used in different ways and in varying degrees. It is beyond the scope of this study to describe the specific objectives of each course. However, all English courses were one-semester non-credit required courses in academic programs focusing on reading and/or writing skills, whereas courses in French and Spanish offered credit.

Data Collection and Procedures. The Student Questionnaire (SQ): To explore the students' perceptions about their e-learning experiences, we used the modified version below of a previously designed instrument [7], consisting of two open-ended questions: *What areas of writing and/or reading in English do you feel you have made progress in, with the use of e-learning? Are you aware of any other outcomes related to your use of e-learning (aside from your improvement in English)?*
YES / NO Whatever your response, please explain.

The SQ was administered toward the end of one semester or after the course ended (spring/summer, 2006). As students were given the option of responding in Hebrew or English, translations were made when necessary. All the students were consenting

participants. To explore teacher perceptions about their e-learning practices, we administered a questionnaire (TQ) consisting of 15 open-ended questions toward the end of the semester or after the course ended. We conducted semi-structured interviews with three teachers who used e-learning extensively in their EAP courses. The same questions were used as in the TQ, but at points during the interviews, spontaneous dialogues between the interviewer and the teacher developed. Interviews were tape-recorded and transcribed. We conducted two interviews with students who volunteered to further reflect on their experiences with e-learning. They were selected on the basis of their responses to the SQ. Notes were taken during the interviews.

Data Analysis. Procedures in descriptive analysis were used for analyzing the responses to questions 1 and 2 of the SQ, where each declared outcome was counted as a response unit. Content analysis was performed on all the responses. Content analysis of the responses to the TQ and the student and teacher interviews was also performed.

3 Findings and Discussion

3.1 Implementation of E-Learning in HE

We present a description of how e-learning is being implemented in EAP courses in Israel based on content analysis of teacher interviews and questionnaires. Most of the universities and many colleges use the same web shell, Highlearn, for web-supported courses. Despite one collaborative effort, the Meital Project, we found much diversity within institutions and among them. For example, only the Open University offers a choice of EAP on-line (distance learning) or in a traditional setting. The other institutions observed in this study offer a variety of EAP courses which include some type of e-learning elements and can be categorized as follows:

Blended Learning: Combination of Online or Distributed Learning and Traditional Face-to-Face (F2F) Instruction [8 Ch.1]. In blended learning courses, students meet teachers regularly, but not exclusively, in a learning center or lab. E-learning is an essential part of the course both during and outside of class. During class, they enter the course website on the university portal and often work in pairs. The two teachers interviewed who have adopted this approach report: "We've been doing it for so long, we can't imagine teaching any other way." They expect their students to continue working after class, asynchronously, and to communicate with each other or with the teacher while doing homework, forums, and projects. Interaction using email and responses in forums is vital to activate writing skills, and they are evaluated on this aspect of learning along with other course components. The communicative ingredient is a vital aspect of e-learning environments which may be a viable alternative to more traditional approaches, especially in large classes which is the trend in HE in Israel because of budget constraints.

Blended Learning in Smart Classrooms. Some courses take place in traditional classrooms equipped with a computer and large screen allowing teachers to integrate e-learning with traditional approaches. Such courses often have their own website, though this is not required as the Internet can be accessed directly. In a graduate

Business English course, for example, different websites are examined synchronously and discussed, activating reading and oral skills. One teacher and her students explore a company website described in an article read by students. Class polls using course websites are held on issues related to articles read for homework. At a conference on "Integrating web-enhanced learning into EFL/ESP courses at the university level" (The Hebrew University, 2006), teachers expressed a desire for more smart classrooms since this setup gives autonomy and flexibility and is less costly than labs.

Traditional Classrooms with Occasional E-learning Experiences. Most of the teachers in this study hold their classes in a traditional F2F learning environment with occasional visits to a learning center or lab where they introduce students to the virtual elements in their course. In the lab at one university, students enter the course website and read selected internet texts as enrichment and background information for texts to be read at home or do grammar exercises and other structured tasks to reinforce classroom learning. Students are encouraged but not required to access the course website at home or during self-access time in the center. However, we found that unless teachers assign specific homework tasks to be done asynchronously, and then relate to them in class, students tend to use the course websites very rarely.

Thus, teachers have much leeway in implementing e-learning activities.

3.2 Student Perceptions

In response to open-ended question 1 of the SQ – What areas of writing and/or reading in English do you feel you have made progress in, with the use of e-learning? 87 students out of 100 described areas of progress. Only 16 perceived no progress; 3 of these reported that they had not used the computer so they could not relate to the question. In response to question 2 – Are you aware of any other outcomes related to your use of e-learning (aside from your improvement in English)? – 73 students out of 100 answered positively; 26 answered negatively; 3 of these students reported that they had not used the computer and 1 did not respond. When asked about the nature of the other outcomes (Whatever your response, please explain.), students described a variety of perceived outcomes gained from their experiences with e-learning.

The Categories. Data reduction and sorting yielded a variety of categories of outcomes derived from the responses to question 1 of the SQ. However, not all of the responses related to aspects of reading/writing in English since some students did not distinguish between what they perceived to have learned in reading/writing skills in English and other outcomes gained from their e-learning experiences. Thus, in Fig. 1 we present the categories relating to improvement in English, while in Fig. 2, we present the categories of other outcomes based on responses derived from both questions 1 and 2. We now present the highlights of the findings:

Improvements in reading/writing in English: As seen in Fig.1, the largest number of response units (23) was in the category "General Reading Improvement" but many of these responses reflected students' awareness of various aspects of reading (e.g., fluency, speed, scanning, understanding of text structure). The categories of "Vocabulary Enrichment" (18) and "Grammar" (18) also revealed students' awareness of different language components. Only 9 responses related to improvement in

Categories from Q1 (Total: 76)	Aspects
General reading improvement (23)	Fluency and speed, scanning/understanding text structure, answering questions, understanding without dictionaries, using reference material to improve understanding
Improved writing (9)	Expressing oneself in English, spelling, formulating sentences, summarizing
Vocabulary enrichment (18)	Using Babylon and other dictionaries, technical English for other fields of study
Grammar (18)	
General English improvement (8)	

Fig. 1. Student perceptions of improvement in reading and writing English in response to Q1

Categories (Total: 162)	Q 1	Q2	Aspects
General knowledge	4	31	Knowledge about the world
Communication, Interpersonal collaboration	11	11	e-mail contact with teacher and students, transmitting information, forums, working in pairs, conversations in class
Independent learning		4	
Real life needs		12	Writing CV, info. about U.S. universities, news, bank accounts, information about university courses
Using the internet as a tool	14	17	Getting information from home, background info, quick access, clear, can hear speech, flexibility of time, place and methods, teacher directed searches
Transfer of skills to other fields	5	5	Other languages, academic texts
Using on-line dictionaries	3	3	
Reinforcement	3	5	Practice, getting homework
Listening comprehension improved	2		
Course information on-line	4		Important for absentees
Allows for feedback	2		
Affective attributes/motivation	8	10	Self-confidence, no fear of criticism, enjoyment, less boring
Help with homework	8		

Fig. 2. Student perceptions of other outcomes derived from Q1 and Q2

writing, implying that students' e-learning experiences often lacked this important skill as expressed by two students: "We need more feedback for writing" and "We had to answer multiple choice questions so we could not improve our writing."

Other outcomes: As seen in Fig. 2 above, the categories relating to other outcomes were more varied and numerous. Furthermore, the total number of positive response units relating to other outcomes (162) by far outnumbered the positive response units (76) relating to reading/writing improvements (Fig.1). This interesting finding implies that e-learning may be affecting our students in more diverse ways than expected. The two largest categories were "General Knowledge" (35) and "Internet as Tool" (31). Students identified many benefits gained from their e-learning experiences as seen from the following quote: "Using the computer you reach all kinds of sites—even if you were not intending to look at them. So you read material you never thought you would read. In addition, it is much easier to learn on the computer because the search is easier and quicker…I learned a lot from the computer even though I didn't mean to learn—and my education improved without meaning to."

In contrast, one of the smallest categories, "Independent Learning" (4) is of special interest since this is one example of many student learning outcomes that are the focus of HE [9]. Also, this is the only category that emerged exclusively from responses to the second question in the SQ. In her interview, a student described how she used e-learning to teach herself German in her spare time: "Reading and listening to the German at the same time helped me become familiar with the words. E-learning helped me to study independently."

An important category we called "Communication/Collaboration" (22) reflected the benefits gained from students communicating with teachers and with each other through e-mail, forums, and working in pairs in class or at the library. One student, a Chinese speaker, described in her interview that her partner in the e-learning sessions helped her by translating technical instructions from Hebrew to English. She also said, "E-learning brings us closer to the teacher; it closes the distance." Another student said, "…email helped me to be in contact with the teacher since I sometimes had questions to ask and clarifications…"

Many responses clustered into a category we called "Affective Attributes" (18). Some students felt that e-learning was a positive experience in general; others expressed a gain of self-confidence, loss of fear of criticism, enjoyment, and the feeling of being less bored while studying. Some of the students took intensive summer courses and were grateful that e-learning made their studies more interesting.

An interesting category "Transfer of Language Skills to other Fields of Study" (10) revealed that students applied what they gained from their e-learning experiences to other areas of study, including their own. For example, one student said that she could deal with academic texts better. Another said, "On the net there are a lot of sources in English in every field, as well as in my specific field of study: nursing. I often use these sources in English—so that I combine 2 goals: improving my knowledge of English as well as learning something connected to my field of study."

The category we called "Real Life Needs" (12) is particularly interesting since the responses related to practical, concrete benefits gained such as learning to write CVs, getting information about U.S. universities and courses, reading the news and others.

The remaining categories consisted of fewer responses but were interesting nonetheless. Each of these categories, "Reinforcement of Learning," "Course Information On-line," "Allows for Feedback," "Using On-line Dictionary" and "Help with Homework," reflects support for learning which students perceive they are getting as a result of their e-learning experiences.

Negative outcomes: When analyzing the negative responses to SQ2, we saw that ten of the students gave no explanation for why they had not perceived any other outcomes as a result of e-learning. However, some said that they were already familiar with e-learning techniques and the Internet and therefore did not learn anything new. Others said they preferred traditional methods, the printed page, books and "a teacher on stage!"

3.3 Teacher Perceptions

Analysis of the data in the teachers' interviews and questionnaires revealed much corroboration with students' reports about their e-learning experiences, but different themes also emerged. In addition, despite their common views on many issues,

teachers did not always share the same approach to e-learning. Due to limitations of space, however, we can only present a few illustrations of these various issues.

Regarding collaboration and communication through e-mails and forums, the role of teacher as guide, and increased interest and efficiency, teachers felt that these were important ingredients in their courses, which corroborated what many students had reported. As one teacher stated, "I think there is a lot more contact with the students…very soon you have a dialogue going which is only possible through the electronic medium…" Another teacher noted that her students "found it exciting knowing we were going to the Learning Center as a class to work on the computer together and do something different." In a teacher training course, students engaged in creative collaboration to design a new teaching tool by posting and getting feedback on the course website.

Teachers also generally agreed that e-learning made academic studies more authentic and relevant, in particular when using the Internet as a tool or resource, which is reminiscent of what many students described as their other outcomes. In the words of two teachers: [e-learning] "Bumps us into the 21st century" and "Takes English learning out of the textbook into the Internet and the real world."

All but one teacher agreed that e-learning should be integrated into the traditional classroom to some extent. One teacher (French for academic purposes) felt that students should engage in e-learning only outside the classroom by doing homework and practicing tests instead of taking up class time. This tells us that some teachers may still maintain that real learning goes on in the classroom in a teacher-centered environment.

One theme that was highly emphasized in teachers' self-reports was the need to encourage more student responsibility for their own learning. Teachers do this by posting homework, announcements and course materials on websites, expecting students to "do the work" even if they have been absent from class. One teacher describes this as her way of extending the classroom so that students play a more active role. Students could ask her questions by email any time of day and get immediate responses so that learning was going on all the time, on a one-to-one basis, not just during group meetings where students were often too shy to speak up or reveal their problems.

Some themes that emerged from the teachers' data dealt with the teacher's responsibility/role and related difficulties. Teachers felt that greater demands are placed on those who work with e-learning. They have to be more flexible and aware of different electronic resources; they need technical ability to maintain websites, develop their materials and train teachers. All this not only leads to more work for the teacher but to occasional resistance from those who prefer not to engage in e-learning.

4 Conclusion

This study has provided a picture, though partial, of how e-learning is being implemented in various institutions of HE in Israel based on teacher and learner perceptions. It has also explored student learning outcomes in the context of e-learning activities in EAP courses as perceived by the learners themselves. One conclusion we can draw is that in the domain of EAP, blended learning is evolving as

a means of enriching prevailing learning processes and making instruction more efficient and flexible. Indeed, the teachers in this study have reported that e-learning allows for more efficient and flexible instruction. They have voiced the view that it is creating more interest in learning among students and a closer relationship between teacher and student, despite the trend toward larger classes.

The learners themselves have implied similar views in their description of the wide range of outcomes they perceived to have gained from e-learning , e.g., reading improvement, vocabulary enrichment, grammar; self-confidence, using the Internet as a tool for studies beyond English, communication and collaboration with fellow students and teachers. These outcomes are examples of cognitive and affective attributes and abilities described by Frye [9 p.6] which can be "a measure of how their experiences have supported their development as individuals." According to Frye, "…student learning outcomes constitute useful measures of quality … They are consistent with the stated missions of higher education." (p.5).

We also see the connection between some of these outcomes and Debski's [4] view of language learning through technology cited in the introduction. Both learners and teachers have acknowledged that e-learning allows for collaboration among students, goal-driven activities and reflection as in the following student response: "I believe that it's very important to use different ways to learn English."

As we look towards making the most of the new technology, we return to Shneiderman's [1] four stages: collect-relate-create-donate. We have witnessed student perceptions that the collection of information is greatly enhanced by the technology. Students have also made clear that communication with teachers and students – relate – is an important use of technology which enhances their learning. Teachers and students have reported on the use of e-learning to create collaborative projects and presentations. Students have found creative ways to transfer their learning to other fields of study; the involved teacher can find innovative ways to use the technology and create a broader learning environment to meet the needs of the 21st century. However, this study implies that the potential for creative use has not been fully realized. To some extent then, we see the first three of Shneiderman's four activities reflected in student and teacher responses. It is the fourth activity, donate, that has still not made an appearance. What use can students make of their new skills to donate to the community in which they live? How can teachers instill in their students this sense of responsible scholarship? Shneiderman sees this goal as an outgrowth of a technology that serves human needs. Further studies on the e-learning tasks themselves would contribute to our understanding of how we might foster such an approach and train teachers to realize the potential that e-learning offers.

We also acknowledge the negative responses toward e-learning expressed by some students. In some cases, they were due to a lack of exposure to e-learning but in other instances, students expressed preferences for traditional methods, the printed page, learning from a book and were also bothered by technical problems. Some teachers also resist e-learning for similar reasons. Individual teaching and learning preferences should be acknowledged and respected; e-learning should not be imposed on everyone. Further research, however, on what specifically creates resistance among some teachers and students might help reduce negative attitudes in the future.

Three dominant themes emerged from this exploratory study: the need for teacher flexibility, the encouragement of learner responsibility, and communication between

teacher and student as well as students among themselves. Thus, for those who wish to become "the guide on the side" and foster collaboration rather than competition in educational settings, e-learning certainly has the potential for achieving these goals.

References

1. Shneiderman, B.: Leonardo's Laptop: Human Needs and the New Computing Technologies. The MIT Press, Cambridge (2002)
2. Shneiderman, B.: Foreword. In: Debski, R., Gassin, J., Smith, M. (eds.): Language Learning through Social Computing. Applied Linguistics Association of Australia & The Horwood Language Center, Melbourne, pp. v–viii (1997)
3. Nachmias, R., Ram, J., Mioduser, D.: The Study of a Campuswide Implementation of Blended Learning in Tel Aviv University. In: Bonk, C.J., Graham, C.R. (eds.) The Handbook of Blended Learning: Global Perspectives, Local Designs, pp. 374–386. Pfeiffer Publishing, San Francisco (2006)
4. Debski, R.: Support of Creativity and Collaboration in the Language Classroom: A New Role for Technology. In: Debski, R., Gassin, J., Smith, M. (eds.) Language Learning through Social Computing. Applied Linguistics Association of Australia & The Horwood Language Center, Melbourne, pp. 39–65 (1997)
5. Cohen, A., Nachmias, R.: A Quantitative Cost Effectiveness Model for Web-Supported Academic Instruction. Internet and Higher Education 9, 81–90 (2006)
6. Strauss, A., Corbin, J.: Basics of Qualitative Research. In: Grounded Theory Procedures and Techniques. Sage, Newbury (1990)
7. Katznelson, H., Perpignan, H., Rubin, B.: What Develops along with the Development of Second Language Writing? Exploring the By-Products. Journal of Second Language Writing 10, 141–159 (2001)
8. Graham, C.R.: Blended Learning Systems: Definition, Current Trends, and Future Directions. In: Bonk, C.J., Graham, C.R. (eds.) The Handbook of Blended Learning: Global Perspectives, Local Designs, pp. 3–21. Pfeiffer Publishing, New York (2006)
9. Frye, R.: Assessment, Accountability and Student Learning Outcomes http:// www.ac.wwu.edu/ dialogue/issue2.html

AudioStoryTeller: Enforcing Blind Children Reading Skills

Jaime Sánchez and Iván Galáz

Department of Computer Science
University of Chile
{jsanchez,igalaz}@dcc.uchile.cl

Abstract. Children tend to learn language conventions through processing environment stimuli. Thus, strategies for reading comprehension are commonly used for this purpose. This paper introduces AudioStoryTeller, a tool for pocketPC to support the development of reading and writing skills in learners with visual disabilities (LWVD) through storytelling, providing diverse evaluation tools to measure those skills. We implemented usability and cognitive evaluation to the AudioStoryTeller software. In the usability evaluation, the easiness of use of the proposed hardware by LWVD was established. The goal of the cognitive evaluation was to measure the development of reading skills through interactive audio narrations using a pocketPC device. Results indicate that users were able to utilize effortless the pocketPC device. AudioStoryTeller software together with cognitive tasks, can contribute to the development of cognitive skills in LWVD. This application allows LVD to have access to unlimited scope of books not available in printed Braille.

Keywords: Learners with visual disabilities, blind children, reading skills, PDA, pocketPC, Text-To-Speech, Audio.

1 Introduction

In the course of their life-span people learn some rules of the language from diverse stimuli, facilitating verbal and written communication and saving from learning them formally. The early use of these rules and their integration are fundamental for the development of the intellectual capacity of these people and their later integration into society [2,18].

Braille is commonly used to read and write by LWVD. One of the main disadvantages is that it is a too slow process compared to graphic sources decreasing the child's motivation and restraining the capacity to develop reading comprehension skills such analysis, abstraction, and evaluation.

Diverse mobile applications have been used for learning purposes [3,4,9,14,15,16,17]. Grams [8] developed a study to determine the motivation for learning by using a PDA, finding that the use of these devices motivated more independent reading than traditional ways of reading. Books embedded in mobile devices such as e-books improved the motivation to read. Freeman [6] predicted that

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 786–795, 2007.
© Springer-Verlag Berlin Heidelberg 2007

this way of book reading will be very popular in the future by integrating them in work spaces and everyday life of ordinary people.

Methodologies [10] and applications [11,12,13] have been developed to support learning in LWVD obtaining positive gains in the development of tempo-spatial skills, memory and mathematical logical analysis. Nevertheless, these applications restrict their implementation to a static context as the school classroom and home. In this study, we want to extend this work to other places and contexts that require users to carry out their readings as sighted people carries them in subway, bus, airplanes and the like, by taking advantage of their dead times. Discovering how technology supports reading activities for LWVD and the issues involved by the loss of vision is a key goal in our investigation.

We present AudioStoryTeller (AST), LWVD-centered design based on Text-to-Speech (TTS) for pocketPC. This software is a tool to support the development of reading and writing skills in LWVD through storytelling, providing diverse evaluation tools to measure these skills. AST has been implemented through auditory interfaces and all users' actions are performed through auditory feedback. A visual support for a fast localization by facilitators when blind children are performing tasks with AST is also provided.

2 AudioStoryTeller

AST is an application to support reading and writing abilities of LWVD through the narration of stories containing several evaluations forms of these skills, such as questionnaires, word replacement by synonymous, construction and completing of new chapters. The user's interaction is through audio interfaces. Each user's action contains a corresponding audio feedback, considering also visual support for children with residual vision. Several design hints proposed by Dunlop and Brewster [5] have been considered. However, considering the target users the multi-task option was removed.

AST is a contribution to the storytelling context and reading software for pocketPC such TextAloud[1] and Mobipocket[2], because it considers virtual access to any text that a teacher or tutor consider relevant. They just need to have a digital document in text format to integrate it to a text list to be reproduced by AST. This tool provides: 1. Mobility to the user, 2. Taking advantage of dead spaces during activities, and 3. Personalized evaluation through automatic evaluation. All these features are intended to improve the autonomy of learners and early stimulation of self-learning techniques.

2.1 Navigation Model

The metaphor utilized is one of a library. The reader chooses a book, selects a desired chapter and then proceeds to read it. The access to different chapters and move within a particular chapter has an important resemblance to natural reading of a sighted individual. When he or she finishes a story accessible evaluations are available.

[1] TextAloud, http://www.nextup.com/
[2] Mobipocket Reader, http://www.mobipocket.com/

AST has four evaluation methods: (A) **Wrong Chapter Selection.** The story to be played has a chapter that does not belong to its dramatic plot, and learners should be able to detect it; (B) **Synonymous.** Several sentences from the story are presented and particular words are to be replaced with a synonym from a list provided by the software; (C) **Questionnaires.** Several multiple choice questions are presented and learners have to choose the right answer; and (D) **Final Construction of the Story**. In this evaluation the child listen to a chapter of a story to be completed, then, he or she can choose the corresponding adjective, verb and noun to each sentence in the chapter (see Figure 1).

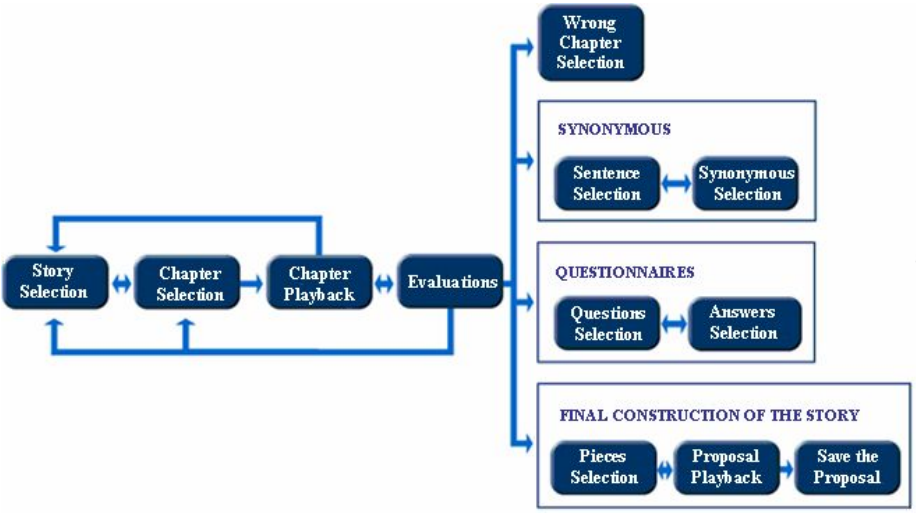


Fig. 1. Navigation Model of AudioStoryTeller

2.2 Control

To interact with AST navigational and functional buttons of a pocketPC are utilized (see Figure 1) in order to provide ample access to different sections of the software, menus and selection of options.

The basic actions a user can perform are search and select. For searching, up and down navigational buttons are used; while for selecting, the central button is used (see Figure 2). Searches can be made upon the stories, chapters, questionnaires, chapter construction sentences, and sentences for synonyms replacement.

During the playback of stories, users can utilize forward, rewind, and pause/resume functions. These functions can be also accessed through navigational buttons, but only in the chapter playback section. To move forward or backwards in a timeframe of 60 seconds, up and down buttons are utilized; while for a 30-seconds timeframe, right and left buttons are utilized. Lastly, the pausing, playing and resuming options are triggered by using the central navigational button.

Current menus are cyclic, so users can select one option easily and quickly, through the central navigational button.

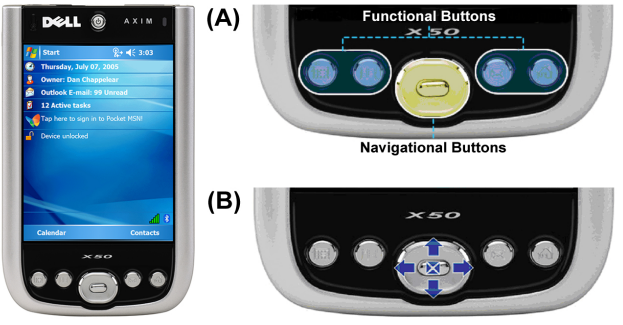


Fig. 2. (A) Functional Buttons and Navigational Buttons (B) Specific Directions of Navigational Buttons (up, down, left, right and center)

2.3 Interfaces: Audio and Visual

End-users of AST are LWVD with ages between 8 and 14 years. For an effective use of the software and considering the target end-users, we used both visual and audio feedback, for users that have residual vision and those totally blind, respectively. Furthermore, the visual display helps the tutor to interact with the software, thus allowing him or her to provide a better guidance for LWVD. Audio cues for the user's actions are implemented. They include: (A) Playing a specific sound when pressing functional and navigational buttons; (B) Reading available items of the selection menu, and (C) Playing the setting information utilized when entering to each section.

The help for available actions and the information for user's awareness are provided through Text-To-Speech (TTS) technology. We avoid mixing TTS and real voice recordings because they can be counterproductive for the proposed goal [7].

3 Usability Evaluation

3.1 Sample

The sample consisted of 11 children with ages between 8 and 14 years old from the school for the blind "Santa Lucía", located in Santiago, Chile. Four of them were boys and 7 girls, 9 with residual vision and 2 with total blindness. The vision disorders were congenital bilateral glaucoma, amaurosis, congenital amblyopia and retinitis pigmentosa. Two special education teachers, experts on vision disorders also participated in the evaluations.

3.2 Instruments

Two instruments were utilized for testing, one to evaluate each usability session and the other to document the observations made to the use of the device by end-users: (A) Summarized End-user Usability Evaluation for PDA, and (B) Non-formal Observation.

3.3 Procedure

Used were tested during seven sessions of one hour and a half each, four sessions in September-November, 2005, and three sessions in July-August, 2006. The whole testing was implemented at the school for the blind "Santa Lucía", in both indoor and outdoor environments.

Each session of the usability testing consisted in the following stages: 1. Software presentation, 2. AST interaction, 3. Non-formal observation, 4. Photographic documentation, 5. Usability evaluation, and 6. Redesign according to the results obtained.

3.4 Results

Learners were able to use the PDA device efficiently and independently. They were able to locate the five lower buttons of the pocketPC without greater difficulties. Children preferred to use headphones instead of the device's speaker, because the volume was too low for them to listen adequately.

The TTS voice utilized was qualified as unappealing during evaluations. Different voices were tested; none of them achieved the desired level of acceptance. Learners mentioned that it could be interesting to add background music or sound effects to the stories, to help them to understand what they were listening and to represent the context of the narration.

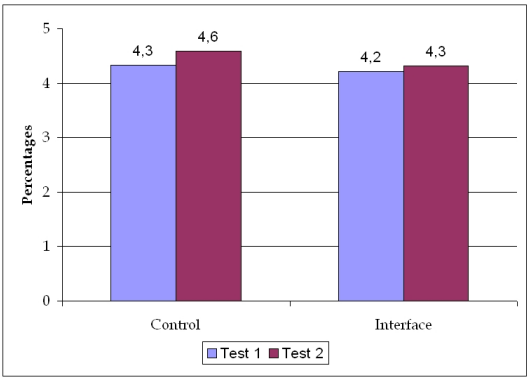


Fig. 3. Usability Results, Test 1 and Test 2

Interaction and interface features were evaluated throughout the development of AST. Interaction was evaluated mainly in terms of the device handling and the interaction with the menus. The data gathered indicate positive results in both tests, observing an improvement of 5.8% between the scores of Test 1 and Test 2. Users evaluated the way the information is presented in the interface and how pleasant was the experience to use it, ending up with a gain of 2.3% between the scores of Test 1 and Test2 (see Figure 3).

4 Cognitive Evaluation

4.1 Sample

The sample consisted of 4 middle school boys with ages between 12 and 14 years old from the school for the blind "Santa Lucía", located in Santiago, Chile. Three of them had residual vision and one was totally blind. The vision disorders they had were congenital bilateral glaucoma, amaurosis, congenital amblyopia and retinitis pigmentosa. All of them were first-time PDA users.

4.2 Instruments

The instrument utilized to evaluate the cognitive impact was the Progressive Linguistic Complexity Test (PLCT) [1]. This test aims to evaluate reading comprehension during the first 8 years of formal education. The test is divided in eight reading levels, each one of them with two different evaluations. The areas evaluated through this test are instrumental area and reading comprehension. On the one hand, the instrumental area evaluates the performance of reading. On the other, the reading comprehension area evaluates how much the learner can retain and understand.

The PLCT is a standardized test for Chile. We used the level 7, A and B schemes, for the pre and posttesting respectively, in macrotype and Braille adaptations. This level was selected because evaluates skills to encounter contextual keys, to place facts in their correct temporal and spatial position, to infer facts based on the global comprehension of a text, and to grasp the intention of the story.

Two cognitive tasks were designed to work with AST. The first task consisted in have children to represent the narrations listened in the software through dramatic group games. In the second task, children had to interpret AST's narrations creating new stories.

4.3 Procedure

The cognitive evaluation setting was placed at the school for the blind "Santa Lucía" during five months, from August to December 2006, comprising 9 sessions of 1.5 hours each. The stages of the evaluation consisted on: (A) Pretesting using PLCT, (B) Developing two cognitive tasks while interacting with AST, and (C) Posttesting using PLCT.

After each cognitive task, learners were evaluated using questionnaires recording specific abilities observed and applied by them.

4.4 Results

The results have been analyzed according to Z-values of the PLCT. The Z-value is a measure unit that let to place a child in comparison with the average and the standard deviation of the group. The ranges are between -3 and 3 points, with an average of 0. By taking into consideration this information the real gain of the students in comparison with their peers was obtained.

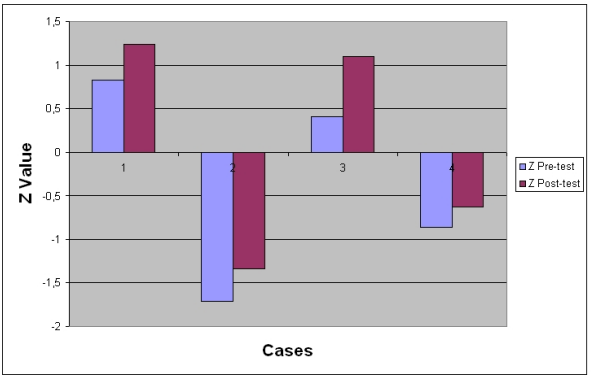


Fig. 4. Comparison of Z value between Pretest and Posttest

Case 3 was a totally blind participant that obtained the most significant improvement compared to the sample. He achieved a score gain of 168%. This result suggests that AST is more suitable for totally blind learners than individuals with residual vision, mainly due to the cognitive models of these users. However, all learners improve their pretest scores obtained. In spite of this progress in their reading skills, two students remain below the average.

In the first cognitive task, the comprehension skills to interpret, memorize and synthesize were evaluated. The total sample obtained a percent average gain of 78,8%, 91,6% y 92,1% respectively. These results support our beliefs that this software help users to identify and assimilate meanings, to remember, to identify and replicate facts, phrases and ideas inside the narrations, and finally, to synthesize what was learned.

In the second cognitive task, the comprehension reading skills enhanced were to analyze, infer, rebuild the meaning and evaluate, the sample obtained a percent average gain of 85,3%, 91,1%, 78,3% y 92,2% respectively.

Moreover, in spite of all the criticism used made about the use of the TTS, the data indicate that it was a very satisfactory and rewarding experience for them to hear their own creations in AudioStoryTeller and share them with their classmates using this media. After this activity was completed, there was a change in their appreciation and valuation of the software, when considering its contribution to allow LWVD to read different texts.

Students were able to take elements from narrations and to shape them into new narrative creations, using their own language and previous experiences.

Finally, by using AudioStoryTeller students were able to use their auditory sense to identify the main ideas in a story, gathering the necessary elements to create a new narration based on their own ideas, interest and knowledge.

5 Conclusions

The purpose of this research was to support reading comprehension in LWVD through interactive narrations in PDA software. Usability and cognitive evaluations provided very interesting and enriching results.

The voice interface used for text reading, a TTS (Text-To-Speech) technology, did not have a good level of acceptance by users. All of them agreed to qualify it as annoying, "robotic-like", plain, and slow, with no emphasis or toning whatsoever, so they suggested adding some mechanism to manage the reading.

Results obtained in this study indicate that there were not limitations for UWVD when utilizing a pocketPC device, since they all agreed that it was comfortable, pleasant to use, and highlighting their shape and size.

Results obtained in the application of cognitive tasks showed that learners were capable of identifying meaningful aspects in the content of narrations and to express them orally, and using body and gesture expressions, through their dramatic representations. To attain this, learners had to interact with the software through comprehensive hearing, analyze texts, identify tempo-logical elements in the narration, and adapt and stage of the play.

It is very likely that this software will help them to identify and assimilate meanings, to remember, identify and replicate facts, phrases and ideas of the narrations, and lastly, to synthesize what was learned.

Students were able to take reference elements gathered through the interaction with AST, in order to assimilate them and create new narrations, using their own language and previous experiences.

Finally, by using AST students were able to provide their hearing sense with an intention, in order to identify the main idea and to capture those elements that allow them to create new narrations based on their own ideas, interests and previous knowledge.

6 Discussion

When individual interpretation embedded in the reconstruction of knowledge are diminished, we can infer that the teaching practices that these LWVD are accustomed to does not promote individual creativity and the free expression of their own ideas during different learning activities engaged at the school.

A PDA device based on audio can reinforce the abilities needed for proper reading comprehension in LWVD. However, cognitive processes through audio and reading comprehension imply the concourse of several skills, so the software should be coupled with other resources.

Learners with visual disabilities can increase their reading comprehension level by using the AudioStoryTeller software as a complementary aid for cognitive tasks and methodological and didactic designs that nurture, extend and widen the learning possibilities of these users, through systematized and continuous use of texts and evaluation resources provided by the software. AudioStoryTeller can be enriched by adding narrations with different degrees of complexity, and evaluations that aim to develop and implement several abilities gradually.

As a result, by using and measuring regularly the achievement and improvement of the level of narrative comprehension and the domain of specific skills for understanding a text, higher levels of reading comprehension can be obtained.

The promising results obtained in this study lead us to state that blind users can utilize this type of software to develop their reading skills, which can be ultimately

become an alternative to Braille books. AudioStoryTeller provides an easy way to widen the story collection; therefore it would be possible to grant access to a greater literature, widening the reading possibilities of learners with visual disabilities.

Acknowledgments. This report was funded by the Chilean National Fund of Science and Technology, Fondecyt, Project 1060797.

References

1. Alliende, F., Condemarín, M., Milicic, N.: CLP Test. Published by Universidad Católica de Chile Editions, Santiago Chile (1991)
2. Allington, R.L.: Critical issues: What´s special about reading program for children who find learning to read difficult? Journal of Reading Behavior 26, 95–115 (1994)
3. Attewell, J.: Mobile technologies and learning: A technology update and m-learning project summary. London: Learning and Skills Development Agency (2005) Available online at http://www.m-learning.org/reports.shtml
4. Csete, J., Wong, Y., Vogel, D.: Mobile devices in and out the classroom. In: Cantoni, L., McLoughlin (eds.) Proceedings of ED-MEDIA 2004 World Conference on Educational Multimedia, Hypermedia & Telecommunications, pp. 4729–4736 (2004)
5. Dunlop, M., Brewster, S.: The Challenge of Mobile Devices for Human Computer Interaction. Personal and Ubiquitous Computing 6, 235–236 (2002)
6. Freeman, M.: Midmorning in the e-book age. Reading Today (2000) http://www.reading.org/punlications/rty/archives/oct_ebook.html
7. Gong, L., Lai, J.: Shall We Mix Synthetic Speech and Human Speech? Impact on Users' Performance, Perception, and Attitude. In: Proceedings of the ACM CHI, March 31 – April 5, 2001, Seattle, Washington, USA, pp. 158–165 (2001)
8. Grams, D.: E-books: Motivating students to read independently. A research project report submitted to Mingyuan Zhang Central Michigan University, EDU 590 (2003)
9. James, J., Beaton, B., Csete, J., Vodel, D.: Mobile educational games. In: Lassner, D., McNaught, C.(eds.) Proceedings of ED-MEDIA 2004 World Conference on Educational Multimedia, Hypermedia & Telecommunications, pp. 801–802 (2003)
10. Sánchez, J., Baloian, N.: Modeling 3D interactive environments for learners with visual disabilities. In: Miesenberger, K., Klaus, J., Zagler, W., Karshmer, A.I. (eds.) ICCHP 2006. LNCS, vol. 4061, pp. 1326–1333. Springer, Heidelberg (2006)
11. Sánchez, J., Flores, H.: Training Blind Children to Develop Mathematics Skills Through Audio. In: Proceedings of the Cybertherapy, Basel, Switzerland, pp. 123–124 (June 6-10, 2005)
12. Sánchez, J., Flores, H., Sáenz, M.: Blind Children Developing Mathematics Skills through Audio. In: Proceedings of ITI 2nd International Digital Media Conference in Egypt, DIGIMEDIA 2005 Cairo, Egipto Marzo 7-8, 2005 pp. 155–166 (2005)
13. Sánchez, J., Sáenz, M.: Sound Immersed Virtual Environments for Blind Children. In: Proceedings of The Fourth International Workshop on Virtual Rehabilitation IWVR '05. Catalina Island, California, USA. September 19-21, 2005 pp. 192–202 (2005)
14. Seppala, P., Alamaki, H.: Mobile learning and mobility in teacher training. In: IEEE International Workshop on Wireless and Mobile Technologies in Education, Växjö, Sweden, pp. 130–135 (2002)

15. Wang, M., Shen, R., Tong, R., Yang, F., Han, P.: Mobile learning with Cellphones and PocketPCs. In: Lau, R.W.H., Li, Q., Cheung, R., Liu, W. (eds.) ICWL 2005. LNCS, vol. 3583, pp. 332–339. Springer, Heidelberg (2005)
16. Williams, M., Jones, O., Fleuriot, C., Wood, L.: Children and emerging wireless technologies: Investigating the potential for spatial practice. In: Proceedings of ACM CHI, Portland, Oregon April 2-7, 2005 pp. 819–828. ACM Press, New York (2005)
17. Wills, S.: Wireless, mobile & handheld: Where are our teachers and students going with their computers? The power of 3. EDUCAUSE in Australia 2001: The power of 3 (Retrieved December 10, 2003) Available online at http://www.gu.edu.au/conference/educause2001/conten2a.html
18. Worthy, J.: Teachers and students suggestions for motivating middle school students to read. Yearbook of the National Reading Conference vol. 49, pp. 441–451 (2000)

Basic Skills Training to Disabled and Adult Learners Through an Accessible e-Learning Platform

Olga C. Santos[1], Jesús G. Boticario[1], Alicia Fernández del Viso[2], Santiago Pérez de la Cámara[2], Carlos Rebate Sánchez[2], and Emmanuelle Gutiérrez y Restrepo[3]

[1] aDeNu Research Group, Artificial Intelligence Department, Computer Science School, UNED, C/Juan del Rosal, 16. 28040 Madrid, Spain
{ocsantos,jgb}@dia.uned.es
http://adenu.ia.uned.es/
[2] SOLUZIONA Accessibility Unit
C/ Acanto 11, Edif. A, 28045 Madrid, Spain
{afernandezde,sperezd,crebate}@soluziona.com
http://www.soluziona.com
[3] SIDAR Foundation-Universal Access.
C/ Sancho Dávila, 35, 28028 Madrid, Spain
emmanuelle@sidar.org
http://www.sidar.org

Abstract. In the paper we present the first evaluation results obtained in the framework of ALPE (Accessible Learning Platform for Europe – EC-029328) a market validation project partially funded by the European Commission through the European Community Programme "eTEN". During the first three months of the project, ALPE service is being customised to the needs of end-users in Spain, the U.K. and Greece. In addition, a small scale test of the service has already been performed by a group of visually and hearing impaired and adult learners coordinated by SIDAR Foundation. In this paper we describe i) ALPE service, including the platform and the courses offered, ii) the validation process and iii) the results from the initial end-user evaluation.

Keywords: Disabled students, Accessibility and evaluation, Basic skills market validation, educational standards, Learning platforms, Open courses.

1 Introduction

The objective of the ALPE service is to improve the human computer interaction (HCI) of disabled and adult learners using an e-Learning platform to access a range of basic skill courses on literacy, numeracy and ICT. For this purpose, the ALPE project has set up an open source, standards based, accessible e-learning platform that delivers courses from a repository of accessible, multilingual, SCORM-compliant courses on basic skills, which draws on the practical lessons learned on pedagogical and organisational issues relating to accessible e-Learning [1]. These courses are to be

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 796–805, 2007.
© Springer-Verlag Berlin Heidelberg 2007

tested in three different countries (Spain, United Kingdom and Greece) by more than 300 students with disabilities. The project is targeted at disabled people (visually and hearing impaired) and adult citizens who want to improve their basic skills, and hence their employability levels.

Basic skills is being defined as "the ability to read, write and speak in the respective language and to use mathematics at a level necessary to function at work and in society in general". This concurs with the Lisbon Memorandum of Understanding (23-24 March 2000) and the Uppsala Swedish Presidency (2002) which reinforced the requirement for training in basic skills such as reading, writing, and arithmetic; lifelong learning; IT proficiency and their fundamental knowledge needed for everyday life.

To ensure the viability of ALPE service, the aim of the project is to assess the service acceptance, referring to the capability of the service to satisfy users' needs in term of the usability, accessibility, affordability and acceptability. In this paper we introduce the platform, the validation process and the usage of the educational content (basic skills curriculum) for the different disability needs of the target groups, including the results from the initial end-user evaluation.

2 ALPE Service

The ALPE service will allow European disabled (visually and hearing impaired) and adult citizens to improve their basic skills. These users will have access to a repository of accessible, SCORM compliant, customised basic skills courses which will be delivered to them via an open source, standards based, accessible e-learning platform that delivers web content education barrier free.

ALPE offers a web portal to the end-users through which they will be able to look for and select the most appropriate basic skill courses for their needs, most of the time with the support of their tutors or mentors. The various courses will be supported through an accessible collaborative e-learning environment. The portal will provide the front-end of the ALPE services, while the back-end will allow for the easy integration of the e-learning platform with existing course modules that follow the SCORM compliant structure, and its subsequent storage in a course repository.

Currently there exists no service intended to support disabled and adult learners in improving his/her basic skills via offering a repository of accessible, reusable courses on basic skills delivered via an accessible platform. There exist some approaches like EducaNext, a multilingual, academic exchange portal, where members of higher education, research organizations, and professional communities can share, retrieve, and reuse learning resources. However, it is not focused on disabled and adult learners needs. On the other hand, there are web-based resources for the assessment and teaching of adult literacy and numeric like the ones produced at BSOLE (Basic Skills On Line for Europe) Project [2], but they are not embedded within an accessible platform addressing the needs of people with disabilities.

ALPE therefore provides a new unique service since it offers access to basic skills courses within an accessible environment, which makes it very attractive for the large

population across Europe of disabled and adult learners who have basic skills problems.

3 ALPE Platform

The ALPE service is provided upon an existing open source learning management system (LMS) called dotLRN [3]. It provides a complete portal framework along with out-of-the-box capabilities for course management, online community tools, content management for publishing in academic settings, and learning management tools for administering tests and tracking results. It features a number of collaborative applications such as calendars, forums, file sharing, assessment tools, etc. Many of this functionality is similar to the one found in other LMS such as WebCT and Blackboard, but it provides additional features made possible by the common data model (e.g. a centralised alert system) and others related to the management of collaborative activities, not just courses. The great advantage here is that this platform is structurally flexible and accessible and supports and accessible delivery of tools and contents based on interoperable standards[1]. Specifically, the following functionality is available for learners with special needs:

- Access to interactive and accessible basic skills courses (SCORM compliant)
- Assessment of the progress in the basic skills addressed in the course using accessible questionnaires defined in the IMS-QTI educational standard
- Grouping users to solve collaboratively a learning activity using accessible collaborative components, such as forums, file storage, comments…
- Scheduling of basic skills course activities through an accessible interface
- Language selection functionality, supporting Assistive Technology available on the client workstation.
- Adapted interface configurations for each learner according to the learners' user model, accessibility requirements and basic skills learning objects.
- Private spaces and information fully accessible through secure interfaces, using secure authentication mechanisms.

The variety of tools provided by the platform support a wide range of interaction situations for users who have different needs. We briefly introduce next a couple of scenarios. On the one hand there is Cristina (a visually impaired learner) that uses assistive technologies (screen reader) to access the information on the browser. The screen reader reads aloud the links and images description. In turn, Cristina interacts with the browser with special (shortcut) keys (alt-number-enter combinations, tab and intro keys). On the other side there are two users, neither of them needing assistive devices. Ifigenia (an adult learner with low vision) just needs to properly configure the browser to see the information written in a large font. On the contrary, Peter (who is deaf) just receives content in either text form or watch videos of a sign language. All the courses are intended to improve each user basic skills. In this way, each user can follow the contents and work on individual and collaborative activities using the functionality provided by the platform.

[1] Furthermore, dotLRN is the first VLE to provide support for IMS LD, levels A, B & C.

4 ALPE Courses

ALPE course repository already provides a set of web-based, SCORM compliant and accessible courses which include interoperability features and cover basic skills on literacy, numeracy and ICT. The courses have been customised from existing courses already developed in the institutions involved in the project, as well as selected from open repositories. These courses are prepared to suit the pedagogical and technical needs for the types of disabilities identified above, and the localization, linguistic and cultural requirements for the three countries involved in the project.

4.1 Methodology for Developing Standard-Based Accessible Courses

A methodology to develop and customise accessible courses has been developed, taking into account the different needs of blind, partial sighted, deaf and hard of hearing people who may not use written language as their first language or have problems understanding standard language, and which take into account the diversity in assistive technologies and converting tools: Braille display, speech output, subtitles, screen magnifications, aural presentation, sign language integration, symbols support integration, access to graphics, … following the requirement of producing easy to 'read' content. This methodology has been agreed and followed by the consortium in order to ensure that the courses uploaded to the platform fulfil the needs of the users, are accessible and standards based.

The following figure (Fig. 1) shows the different processes involved. Since traditionally in the universities courses are prepared in Microsoft Office Word (this is the situation we came up with), we looked for a plug-in that made easier the customisation process to the professors. CourseGenie tool was chosen according to the results of a state-of-the-art done on editors supporting accessible and educational standard-based courses [4]. CourseGenie provides different predefined styles to give the appropriate format to the different elements of the text. QTI Lite questionnaires can be included in the course and Metadata information added. It provides also some automatic validation tests on very commonly used elements, such as images (alt and long description) and tables (row and column headers). Moreover, this Word document can generate SCORM packages, which contain the corresponding XML manifiest with the course items and related resources. In particular, these resources consist on the contents of the Word document separated into accessible HTML pages. In order to improve the accessibility of the output, a Word template with more detailed instructions for course authors and which gathers best practices in accessible documents writing, is also available for authors.

Nevertheless, the process is not concluded with the CourseGenie output in terms of a SCORM package, but several evaluation tasks are performed afterwards in order to guarantee the real accessibility of the contents. First, a syntax validation (HTML and CSS) is performed. Next, there is an external accessibility evaluation, which include both automated evaluation tools and an in depth analysis of the source code and tests with different browsers and configurations. If the course passes these two evaluations, there is some guarantee that the courses are quite accessible. However, a final evaluation with technical aids is performed to guarantee the access to the contents by all users.

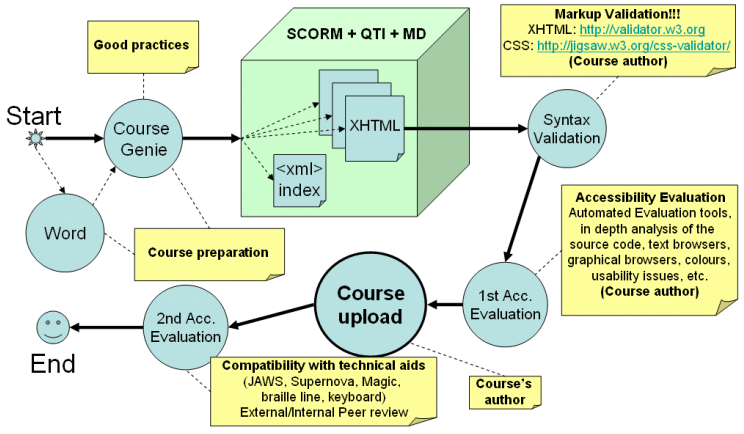


Fig. 1. Process to customise existing courses for ALPE service, resulting in accessible SCORM compliant courses

This initial methodology depends very much on a proprietary text editor (i.e. MS Word). Therefore, it is being extended to cover a wide spectrum of editors and focused on the usage of open source tools. Detailed information is available in [5].

4.2 Adaptation of the e-Learning Courses to Suit End User's Needs

In order to produce accessible contents, the following recommendations have been taken into account. For persons with sensory totally visual disability, besides presenting them the contents in computer legible support for a screen reader which allows them a fluid reading and the use of the transcriptor Braille, the platform offers them some other alternatives. Those learners that have low vision can access the content by applying the appropriate values in the style sheet used, making modifications in the format of the content. In this way, content presentation can be adapted to their needs, offering a suitable contrast of colours or a magnified size of characters that allows an accurate reading. Likewise, both for totally and partially visual disabled, the contents of iconic nature (images, maps, etc) are presented with a sonorous support or a textual description that contributes the required information of all details that the individual is unable to capture. It is also necessary to avoid any type of acoustic pollution. To take into account the difficulties of concentration that these learners can present, it is important to facilitate them in advance any type of material that will allow them to have a mental scheme or general orientation on the matter that is going to be explained.

For learners with sensory auditory disability the channel across which they gather the information is predominantly the visual one. So they have to be provided with all the possible visual helps in order that content is not lost. The information facilitated by means of sonorous signs that the learner will not catch (totally deaf) or will not do it entirely (partially deaf), is supported with visual signs in panels or similar elements that result in an easy visual access by the learner. It is important to guarantee that the learner pays attention to them. With regard to the linguistic difficulties that, due to

their disability learners could have, the use of foreign words and technical terms is limited, and depending on the course, a dictionary of synonymous can be provided to overcome the difficulties of comprehension and expression across written language. Moreover, any instruction about the accomplishment of the activities is facilitated in written form. The pedagogic contents allows to choose an educational modality that leans more on visual contents than in written, simultaneous tasks in which the student should perform two activities at the same time are never presented. Moreover, since the persons with auditory deficits, in general present a few characteristics that limit their social interaction, we try to avoid isolation by finding a form of common communication, encouraging them to take part definitively actively in meetings, gatherings, conferences, forums, etc.

5 Market Validation Process

The goal of the ALPE service is twofold. First, identify individual needs to prove that the services offered are able to attend those particular needs. Second, validate services to all, considering the specific needs of every individual. So first, we will try to identify what type of needs they have, and afterwards, validate whether they have been fulfilled. In order to market validate the service, the project has started with the recruitment of a significant amount of end-users, consisting of visually impaired and hearing impaired students, adult learners and main stakeholders/customers from Spain, the United Kingdom, and Greece. The pilots are conducted in Spain, the United Kingdom and Greece and are undertaken by the following organisations: UNED (the Spanish National University for Distance Education) and SOLUZIONA in Spain, UKOU (the Open University) in the UK and PAB (Panhellenic Association of the Blind) in Greece.

The research methodology adopts both qualitative and quantitative methods, using a variety of research tools including questionnaires (online), workshops, interviews with stakeholders, tutors and customers, and interviews with learners. In Spain, the users for the trials are being selected among 1) the learners with disabilities enrolled at UNED who are experiencing some problems in writing reports or essays and who are highly likely to be in their first year of undergraduate programme, 2) external learners with disabilities who need to improve their basic skills and are enrolled at UNED at the different types of continuing education courses, and 3) potential students for continuous education students that may be aware of ALPE project via end-user organizations such as SIDAR. In particular, in this paper we present the results of a small-scale validation conducted by users with disabilities recruited by the SIDAR Foundation-Universal Access.

The validation process consists on two assessments, pre and post basic skills assessments. At the pre-assessment, learners go through two steps: 1) enrolment and 2) screening. At the enrolment stage, disabled learners are given a questionnaire to be filled in with data regarding their disability and whether the student works or not. In particular, it includes information about the type of disability (physical, psychical, hearing and visual), a textual description of the disability and percentage of disability. The second step consists on testing weather the learner is at an acceptable level of competence in basic skills. This test usually takes 20 to 30 minutes to do and the score

will say ‘you are at level 2’ or ‘ you are not at level 2’, where level 2 is usually the level of competency at leaving school.

Regarding the post-assessment, there are two types. The first one, which is more important from a market validation point of view, consists on a questionnaire and interviews taking place at the end of the market validation with end-users to measure a) attitudes to learning using the ALPE service, b) motivation to improve basic skills, c) perception of their own improvement, d) what their basic skill tutors think of their progress, e) approach to learning using this technology. The second type of post assessment consists on asking the learner to take the screening test again (with a different set of questions). However, there are limitations to this way of measurement since it is never intended to be a purely experimental set-up and the results can never be read as conclusive. What it does is that it gives us an indication of early results/outcomes and the possibilities of conducting a purely experimental set-up with control groups, etc.

6 End User Evaluations

Previous to the evaluation with end-users became a revision of the content generated by the two tools that are being used like mediators between the original course created by the professor in Word and the conversion output to be introduced in the platform, detecting some failures of accessibility in this generated content. For that reason, it was needed "to clean" the contents generated for the test course, being aware that it will be necessary to work on the mediating systems to guarantee the accessibility of the contents.

The objective of the evaluation with end users is to detect the possible failures of accessibility and adaptability that the platform and the contents generated could still have as well as the degree of satisfaction of the user. The end users evaluation was made with the participation of users with visual deficiency, auditory deficiency and adults, in agreement with the objectives of the ALPE project. Participants in the evaluation included users of technical aids and diverse navigators, as well as users with diverse level of experience in using the Internet and e-learning platforms. There were users who had no previous experience in e-Learning and others who had used different platforms.

6.1 Methodology

For the end user evaluation a course with contents on web resources and controlled characteristics was prepared, so that the reception and interaction on the part of the users could be proven. The course had its pedagogical objectives clearly outlined. Moreover, it also intended which users’ needs were covered and which not.

The course was extracted from an existing course on ‘Teaching to Learn through Internet’ being run at UNED from 1999 by some members of aDeNu group under the on-going education program of this university. The contents prepared for this course consisted on the usage of Internet resources, the Google use case and the Web 2.0. It contains textual content (e.g. plain text, titles and subtitles, internal and external links, numbered lists, definition lists, etc.), graphics, images, subtitled video, forms, etc.

Two questionnaires were prepared to be filled in before and after the accomplishment of the course. In this way, and after the analysis of the user interaction traces during the execution of the course and the analysis of their answers to the evaluation questionnaires, we detected successes and failures regarding the navigation in the platform and the understanding of the course contents.

The questionnaire consisted on a first part (previous to the evaluation) to gather their personal data, their degree of visual, auditory or another disability type, their usage of technical aids, their experience in the use of Internet, the inexperience in the use of e-learning platforms, and their disposition towards the use of these platforms. The objective of the second part (to be answered after the course execution) was to detect if the users had found some accessibility failures due to i) the platform, ii) the contents or iii) the user low experience in using the Internet, his/her own browser and/or the technical aids used.

With respect to the accessibility, since end users are not suppose to know the Web Content Accessibility Guidelines of W3C-WAI (WCAG) (V1.0 or the upcoming 2.0) [6] neither to understand the somewhat dark language in which their checkpoints are written, we came to include in the questionnaire the questions on the twelve more common failures (according to what it is deduced from the recent studies ordered by the Disability Right Commission (DRC) of the United Kingdom and by the United Nations)[2], but written up in a clear language for the users. Moreover, the questions focused on emphasising the user perception on the contents and not as much in their technical aspects. For example, as far as the WCAG 7.3 checkpoint says: "7.3 Until user agents allow users to freeze moving content, avoid movement in pages". In this example, what it is requested for the end user is weather there were texts in movement that distracted them and made difficult to them paid attention to the main content.

Moreover, the questionnaire included also questions on the pedagogical objectives of the course and on the satisfaction or dissatisfaction in the usage of the platform.

6.2 First Outcomes from the Evaluation and Future Tasks

The results of this first analysis have confirmed us the objectives expressed in section 4. There has been no surprise as far as the accessibility requirements or the users' perception on their deficiency, when existing. It has been especially significant the outcome regarding the necessity to adapt the language used and to offer glossaries that clarify the terms used within a course.

All the users agreed that it would be desirable that it stands out clearly in the course what information is most important or what information will have to be remembered. Therefore, it is advisable to provide the authors with the support for highlighting reinforcement contents for the students.

Regarding adult users and low experienced in the usage of Internet (who are not familiarized with the options that offer them the browser), we have detected that they ignore the possibility of accessing the meaning of abbreviations and acronyms, provided that they are correctly tagged and that they contain the attribute "title" that

[2] The Disability Right Commission of the United Kingdom report:: www.drc-gb.org/PDF/2.pdf. The United Nations (UN) report: http://www.nomensa.com/resources/research/united-nations-global-audit-of-accessibility.html

expands its meaning. Therefore, for users that ignore functionalities of their user agent it is convenient to offer basic information on the Internet and the accessibility options that their user agent offers. Likewise, it was emphasized the necessity that links to external resources does not make them get lost in the cyberspace. For example, if students are taken to an external site in which "a concept is clarified", that site must present the information in a clear and direct form, so that students do not get lost while reading the new website. Nevertheless, in spite of these comments, all the students agreed in their satisfaction on the level of learning reached and its general valuation was positive.

As previously mentioned, the obtained results to date have not been in any way surprising. In turn, they have allowed us to confirm the appropriateness of the project objectives and the extreme importance of the tasks that are being carried out so far. In the evolution of the project it will be necessary, therefore, to arbitrate the different anticipated options of adaptability, and to maintain guides for the end users of the platform in its two roles: professors and students. Authors are to be provided with a guide that explains them how to take advantage of all the resources that the platform offers and help them to provide the proper support to students at the time of using the platform. Learners are also to be provided with a similar guide on how to navigate through the platform and how to use their user agents, along with the available technical assistance. Our objective here is twofold: (*i*) we try to leverage users' knowledge and experience in taking advantage of platform resources, and (*ii*) we aim at increasing the basic skills of users in using assistive tools.

7 Conclusions and Future Works

ALPE project aims at validating the market of an open source, standards based, accessible e-learning platform through the implementation of innovative and accessible human-computer interaction techniques to provide training to disabled and adult students that enable them to improve the fundamental knowledge needed for everyday life.

This paper has shown several related issues: the platform, the validation process and the usage of the educational content (basic skills curriculum), including the results from the initial end-user evaluation previous to the demonstration milestone defined at the third month of the project (end of March). The demonstration will provide evidence on diverse issues: accessible content production, attending vision and hearing problems (including navigation support), usage of assistive technologies and types of access to learning materials, converting functionalities and services (automatic, supported, manual), and accessibility and usability features of all the provided services.

ALPE service tries to meet the needs of their target learners by providing the following key features: programmatic learning and the presentation of content to suit the learner's level of disability, alternative learning options, additional information to those who need assistance, and peer-to-peer learning when advisable.

The initial end-user evaluation described in this paper provides the first input for the market validation to be performed for ALPE service, and it will be framed within a series of iterative evaluations, which are intended to identify improvement issues.

Apart from the users recruited at UNED as described in the paper, in the UK, learners are being drawn from further education (post-16) colleges or Open University Access or Foundation Level (equivalent to preparatory or first year undergraduate) courses and in Greece, the learners are being selected among members of non-profit organizations that support visual and hearing impaired adults. This large-scale validation will start in March 2007. In total, 300 users will be involved in that evaluation, 100 from each country.

Moreover, ALPE project will also benefit from the Computer Training Program that PAB will undertake with more than 700 visually impaired users. This program is funded by the Greek government and aims to give a certified training in Computer Basic skills to 750 blind and partially sighted Greek people (with up to 50% degree of disability) both employed and unemployed, from 16 to 65 years old.

Acknowledgments. The authors would like to thank the European Commission for partly funding ALPE project (eTEN 029328). We would also like to thank SIDAR Foundation, and among end-users, special thanks to Martín Baldassare, Pedro Novoa and Enrique Gutiérrez for their useful comments from the evaluation.

References

1. Cooper, M., Colwell, C., Taylor, M.: Practical Accessibility in eLearning: Decisions, Dilemmas and Discussions for Educators and Developers. In: Workshop on Practical Accessibility in eLearning as part of the Accessible Design in the Digital World conference, August 23-25 Dundee, Scotland (2005)
2. BSOLE: Basic Skills On Line for Europe, eTEN Project under contract C28018. D3: Market Information, Culture and Pedagogy: A report on the market and a look at customisation and adaptation issues in the translation of Basic Skills On Line for Europe (2003)
3. Santos, O.C., Boticario, J.G., Raffenne, E., Pastor, R.: Why using dotLRN? UNED use cases. In: 1st International Conference on FLOSS: Free/Libre/Open Source Systems (2007) (in press)
4. Santos, O.C.: Technology Enhanced Life Long eLearning for All. In: Maillet, K., Klamma, R.(eds.) Proceedings for the 1st Doctoral Consortium on Technology Enhanced Learning. European Conference on Technology Enhanced Learning, p. 66–71 (2006)
5. Santos, O.C., Boticario, J.G., Rodríguez Ascaso, A., Gutiérrez y Restrepo, E., Barrera, C.: Accesible and Reusable courses on ALPE Platform (available in Spanish). In: 1st International Conference on FLOSS: Free/Libre/Open Source Systems (2007) (in press)
6. WAI Content Guidelines for creating accessible web pages (WCAG) v1.0, available at: http://www.w3.org/TR/WAI-WEBCONTENT/ WCAG v2.0 (draft at time of writing) available at: http://www.w3.org/TR/WCAG20/

Interface of Online Mini-Go-Game with Pen Input Guide for the Blind

Michio Shimizu[1], Masahiko Sugimoto[2], and Kazunori Itoh[3]

[1] Nagano-ken College, Miwa 8-49-7
Nagano-shi, 380-8525 Japan.
michio@nagano-kentan.ac.jp
[2] Takushoku University Hokkaido College, Memu 4558
Fukagawa-shi, 074-8585 Japan.
sugimoto@takushoku-hc.ac.jp
[3] Faculty of Engineering, Shinshu University, Wakasato 4-17-1
Nagano-shi, 380-8553 Japan.
itoh@cs.shinshu-u.ac.jp

Abstract. In this paper, an interface of a Mini-Go-game (Go game with a 9x9 board) on a network for the blind is studied. If the Mini-Go-game on the Internet for the blind can be realized as a web application using PC or a cellular phone, it is easy to obtain an opponent without going out. Therefore, we introduce a pen input guide which correspond to a 9x9 board, and propose a co-operative system of a Mini-Go-game with sound feedback.

Keywords: Mini-Go-game, visually impaired, pen input guide, sound feedback, co-operative system, virtual sound screen.

1 Introduction

The use of the Internet has been considering as the communication support to a visually impaired person in order to increase his life quality. In this paper, an interface of a Mini-Go-game (Go game with a 9x9 board) on a network for the blind is studied. If the Mini-Go-game on the Internet for the blind can be realized as a web application using PC or a cellular phone, it is easy to obtain an opponent without going out. Therefore, we introduce a pen input guide which correspond to a 9x9 board, and propose a co-operative system of a Mini-Go-game with sound feedback. Our system consists of a server, two clients and the input-output interface uses tactile and auditory senses together.

2 Background of the Mini-Go-Game

Go is one of a board game for two players. It is also called Weiqi in China, Igo in Japan, and Baduk in Korea [1]. The game of Go is now popular throughout the world especially in the East Asia. However, for a visually impaired person, the Go game of a normal 19x19 board is too difficult to play. Although the game rules are very

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 806–812, 2007.
© Springer-Verlag Berlin Heidelberg 2007

simple, it is very difficult to grasp the layout on a board. Then, using the 9x9 board game of Go (we call it Mini-Go-game) (Fig.1), the special equipment is made for the blind (Fig.2) [2]. The go board has holes to keep stones and black and white stones can be identified with a mark on them. Touching lines and stones with finger enables the visually impaired to position a stone and imagine a stone layout. The player in a game reads aloud the position of a stone each other.

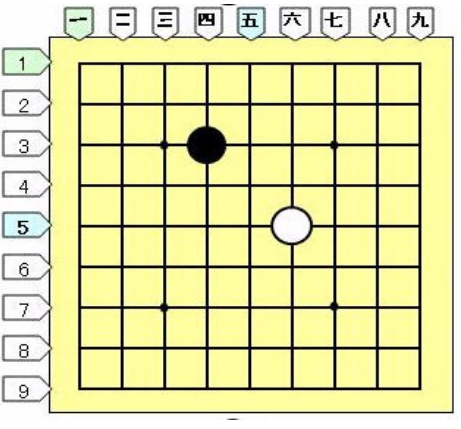


Fig. 1. Mini-Go-Game

Fig. 2. Mini-Go-game board for the blind

By the way, the relation of Go and the Internet & computer is becoming very strong. Recently, many (sighted) peoples are playing the Go game (in this case a normal 19x19 board) using the Internet against opponents from around the world. It is simply called “Internet Go”. On the other hand, “Computer Go” is the field of an artificial intelligence [7]. In spite of great efforts to develop a program, it is

considered that the Computer Go is still infant level. As for the "Computer Mini-Go", very strong programs were created and they are helpful to study the Go game.

However, the interface of "Computer Mini-Go" and "Internet Mini-Go" for a blind person is not realized. Only so-called "Mail Mini-Go" is performed by using the e-mail among blind persons. The player sends the position of a stone as a numerical value through e-mail, and the opponent reads out it by using a screen reader and sets a stone on his real board. The system played on the Internet while grasping the layout of a stone on the 9x9 board is not developed yet.

3 Co-operative System Using Pen Input Guide

Then, we create the virtual board on web corresponding to the real 9x9 board, and make the co-operative system (Fig.3) with which the blind can play a Mini-Go as an experiment. The system is made from a web server and two clients with a headphone and a pen tablet. We introduce the pen input guide (Fig. 4) [3] on a pen tablet to grasp a layout of the stone on a board with sound feedback. Sound feedback is made from a virtual sound screen (Fig.6) [4] that is a 2-dimensional space using characteristics of hearing, called sound localization.

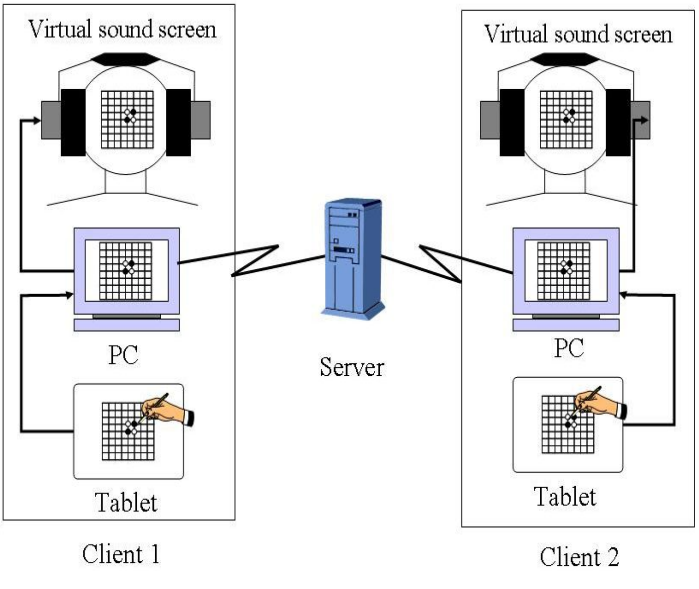


Fig. 3. Co-operative system

3.1 Pen Input Guide

Our pen input guide is made from acrylics with thickness of 4mm, and is used on a pen tablet (Fig.4). The guide consists of a game board area and eight operational holes to the surroundings. Game board area (Fig. 5) is 54 by 54mm in square, and has 81

holes located in a line in the shape of a lattice, and slots for making movement of a pen easy. Among 81 holes, five holes are enlarged and these holes are equivalent to the center point and the star points of a real board. Furthermore, in order to make a game board area clearly, the edge on all sides is high.

On the other hand, operational holes have functions for supporting advance of a game. Eight operational holes express the function of “start”, “pass”, “send”, “get”, “count”, ”arrange”, “touryou”(means give up), and “history” to a clockwise rotation from the upper left. Placing a stone or calling a function is performed by clicking the side button of a pen, pointing out the hole with a pen.

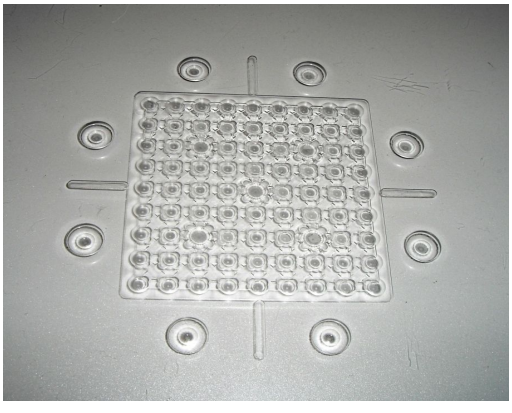


Fig. 4. Pen input guide

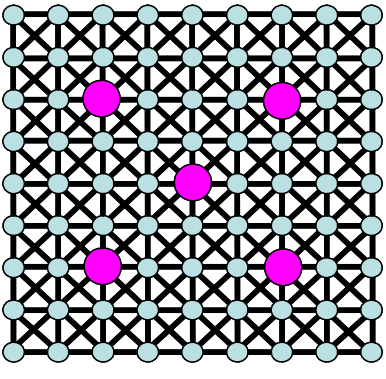


Fig. 5. Game board area

3.2 Sound Feedback

In order to check the layout of stones, three kinds of tones corresponding to black stones, white stones and vacant intersections are assigned in advance. The musical instrument sounds of a piano and a violin are assigned to black and white respectively, the sine wave is assigned to vacant. Moreover, a property of localized

sound which has a position-information specified uniquely on the 2-dimensional virtual sound screen is added to these three kinds of tones.

Figure 6 illustrates 81 discrete localized sound positions using by our system. The vertical axis of the virtual sound screen is expressed by 9 frequencies of the musical scale G4 (392.00Hz) through B5 (987.77Hz). The horizontal axis is expressed by 9 interaural level differences between both ears. The difference of two adjacent localized sounds set at 5 dB. This creates a virtual sound screen made from 81 discrete localized sound positions.

If the hole on a board area is followed with a pen, one of three kinds of sounds can be heard. The time duration of the sound output is 30 msec. Other feedbacks using speech synthesized sounds which support the advance of a game are as follows.

(a) Reading out the operational function when a pen is on an operational hole.
(b) Reading out the coordinate when a stone is put on a vacant intersection.
(c) Reading out the sequential coordinates by the function “history” in order to check the game history from a start to present.
(d) If he is going to input a stone at the place on which a stone is already put, a sound "sokoniwa-utemasen" (means you can not input it there) is outputted and he can not input.

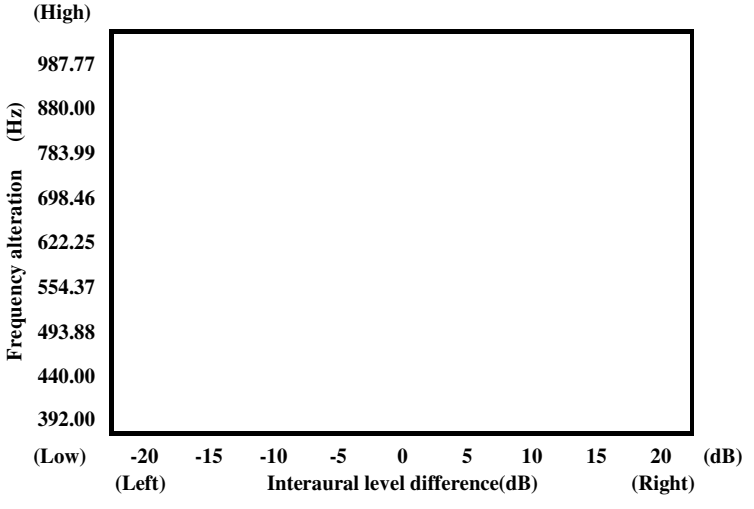


Fig. 6. Virtual sound screen

3.3 Game Program

Let suppose player P1 and P2 are going to start the game, and player P1 and P2 use client PC C1 and C2 respectively. A flow of our game program consists of (1) game start, (2) stone placing and (3) game end. A player cannot be input a stone at the time of the turn of his opponent.

(1) Both players' input of “start” starts a game. The player inputted first becomes black. Now, we assume that player P1 became black.
(2) a) Player P1 inputs a black stone into a vacant hole, and the sound corresponding coordinate is outputted. For example, the black stone shown in Fig.1 is “yon no san” (4 and 3). Next, player P1 inputs a function “send”, so the coordinate is sent to the web server.
   b) The web server sends a signal to both clients, and a sound “ping-pong” is outputted from both clients.
   c) Player P2 inputs a function “get”, the coordinate is sent to the client C2. Then the sound corresponding coordinate is outputted.
   d) Player P2 inputs a white stone into a vacant hole, and sound corresponding coordinate is outputted. For example, the white stone shown in Fig.1 is “roku no go” (6 and 5). Next, player P2 inputs a function “send”, so the coordinate is send to the web server.
   e) The web server sends a signal to both clients, and a sound “ping-pong” is outputted from both clients.
   f) Player P1 inputs a function “get”, the coordinate is sent to the client C1. Then the sound corresponding coordinate is outputted.
   g) From a) to f) is repeated.
(3) When one of players inputs a function “touryou” or both of players input a function “pass”, the flow of the game will go to the game end. Each territory is counted by the continuing functions “arrange” and “count”.

Our game program is a developmental stage. The program that players input a stone by turns and they exchange the place of a stone mutually is already created. But, in order to decide a victory or defeat, the program which counts both territories is not made yet. Various board programs are already proposed [7] as a base of “Computer Go”, it is more efficient to use these free programs.

4 Conclusion

It was observed through the practice of the subjects that the use of a finger was important. Almost subjects follow the holes on board area with the left index finger firstly, and next they point out the hole with a pen and input a stone. In this report, the board game using a personal digital assistant was assumed, so it was considered that an input area was comparatively small. However, when performing a game on an ordinary tablet, it is desirable to make an input guide and its holes large. In order to use a tactile sense more, a finger input equipped with the function of a pen input might be considered.

From the comments of the subjects, we recognize that the sound feedbacks of our localized sound are effective to grasp a layout, but the quantitative evaluation of the effects has not been studied. And, the improvement of the method to grasp a layout on a board is still a future work. For example, the virtual sound screen made from HRTF [8] would be applied to read out of the coordinates.

Acknowledgement. We are thankful to Mr. Takayoshi Baba of the Shinshu University Graduate School who cooperated in creation of the input guide.

References

1. http://en.wikipedia.org/wiki/Go_(board_game)
2. http://www7.ocn.ne.jp/s̃go/index.html
3. Masahiko, S., Kazunori, I., Michio, S.: Simple pattern creation using a pen input guide and sound localization for the visually impaired. In: Miesenberger, K., Klaus, J., Zagler, W., Burger, D. (eds.) ICCHP 2004. LNCS, vol. 3118, pp. 745–752. Springer, Heidelberg (2004)
4. Michio, S., Kazunori, I., Tatsuo, N.: Pattern Representation System using Movement Sense of Localized Sound. In: Proceedings of HCI, pp. 990–994 (1999)
5. Roth, P., Petrucci, L., Assimacopoulos, A., Pun, T.: From Dots to Shapes: an auditory haptic game platform for teaching geometry to blind pupils. In: ICCHP 2000 proceedings pp. 603–610 (2000)
6. Kamel, H.M., Roth, P., Shiha, R.R.: Graphics and User's Exploration via Simple sonics(GUESS): Providing Interrelational Representation of Objects in a Non-visual Environment. In: Proceedings of the 2001 Inter. Conf. on Auditory Display pp. 261–266 (2001)
7. http://www.computer-go.jp/index.html
8. Kazunori, I., Michio, S.: GUI Objects Represented by New Localized Sounds using HRTF. In: Proceedings of HCI, pp. 1285–1289 (2003)

Enhancing Universal Access – EEG Based Learnability Assessment

Christian Stickel[1], Josef Fink[2], and Andreas Holzinger[3]

[1] VirtualTrends Int. Ltd. & Co. KG, R&D, Lerchenweg 3, 36137 Grossenlueder, Germany
christian.stickel@virtual-trends.de
2 University of Applied Sciences Frankfurt, FB2 Computer Science,
Nibelungenplatz 1, 60318 Frankfurt, Germany
jfink@fb2.fh-frankfurt.de
[3]Medical University Graz, Institute of Medical Informatics Statistics and Documentation,
Auenbruggerplatz 2/V, 8036 Graz, Austria
andreas.holzinger@meduni-graz.at

Abstract. This paper presents initial research on a new learnability assessment methodology. We propose the use of electroencephalography (EEG) to further improve usability testing. We discovered whether and to what extend there is a correlation between brainwave patterns and the learnability of the software used. Our central hypothesis is that learnability can be assessed by analyzing the rise and fall of specific frequency bands in electroencephalographic recordings. In order to collect empirical evidence for our hypothesis, we conducted an experiment with N=32 participants. We developed a test environment comprising a low-cost EEG system and developed software for analysis and testing. Based on our findings, we consider our EEG-based learnability test applicable, either as a pre-test – in order to determine whether further testing is necessary – or as an augmenting method during standard usability testing. The users' emotions, registered on the EEG, can be applied as a baseline for detecting possible usability difficulties and employed in the development of a biological rapid-usability method for accessibility assessment.

Keywords: Learnability, Biological usability testing, EEG, evaluation method.

1 Introduction and Motivation for Research

Within our information overloaded world, the need for universal access to information is undisputed. In order to assure progress towards this objective, universal test methods have to be developed that can be applied to the information society *without the need for additional adaptations or specialised (re-)design* [1], [2]. *Learnability* denotes the ease with which users can employ a particular piece of software by observing and using it. The main question is: How long does it take a user to learn the handling of the software in order to efficiently achieve a particular goal? Learnability is especially relevant for designing and deploying complex software, mainly due to potential reductions in learning and training time.

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 813–822, 2007.
© Springer-Verlag Berlin Heidelberg 2007

2 Theoretical Background: Brainwaves, Arousal, Relaxation

Everything humans perceive is processed as electrochemical signals in the human brain [3]. These signals are rhythmic, consequently called *brainwaves* and can be measured on the scalp by using electroencephalography (EEG). Brainwaves are always present as different types of sinusoidal EEG activity, changing when the mental and physical state of the subject changes [4]. In order to get more detailed information on brainwaves, their spectrum has to be divided into particular frequency ranges. EEG allows the non invasive observation of electrical processes in the cerebral cortex, which is believed to be largely responsible for our individual thoughts, emotions and behavior. Cortical processes involve electrical signals. Any of the innumerable electrochemical discharges of the neurons in the cerebral cortex produces a small electromagnetic field with a frequency between 1 and 40 oscillations per second. The sum of these signals forms the aforementioned brainwaves, which can be recorded by EEG scanning equipment. Recording is obtained by placing electrodes on the scalp, thereby following a reliable and reproducible system of placement. Each electrode records electrical activity from potentials that contain approximately 30–500 million neurons [5]. The activity of a single neuron can not be interpreted as something intellectual or cognitive. Cognition, intellect, and awareness are bound to global activity states of the brain, so-called macrostates [6].

Similar to all periodic signals, brainwaves can be described by both frequency and amplitude. The frequency is the number of times a periodic event repeats within a second. Widely used EEG frequency bands include gamma (30-40 Hz), beta (12.5-28 Hz), alpha (8-12.5 Hz), theta (4-8 Hz), and delta (1-4 Hz). Frequency ranges of these bands overlap by at least 0.5 Hz. The amplitude represents the power of electrical impulses generated by the brain; its intensity is measured in microvolt (μV) [5].

As already mentioned, brainwave activity changes when the mental and physical state of a subject changes. When a person is aroused, excited, or alert, for example, the EEG pattern has a low amplitude and a fast frequency. When a person is calm and relaxing quietly, however, especially with her eyes closed, brainwaves with a high amplitude and a low frequency can be observed [7].

The ability to mobilize metabolic energy in order to meet environmental or internal demands on behavior is called an arousal response. Three types can be identified based on the duration of a response: alertness = short term, vigilance = long term, and arousal = longer term [8], [9]. Arousal has a direct impact on motivation [10], [11], [12], [13], [14]. Arousal responses of novice computer users, for example, are most likely associated with frustration and discomfort. Misunderstandings of the navigation, error messages, and crashes of the operating system may gradually change the duration and power of the arousal response. With ongoing training and help, however, users become more acquainted, thereby increasing their tolerance towards frustration. In this case, major developmental changes in the arousal response are the growth of processes that control and shape the response. Moreover, the range of internal and environmental stimuli that activate these processes is increased, which can be observed in the context of experienced users (e.g., they remain unaffected of certain types of system errors). EEG can be used for anticipating alertness, vigilance, and arousal. The type of electrical activity recorded can be associated with behavioral arousal; rhythmic fast activity (beta/gamma) appears to be involved in states of high

arousal (e.g., learning, problem solving), whereas rhythmic slow activity (alpha) can be associated with a relaxed, yet alert state of consciousness. Behavior as well as EEG patterns tend to gradually change over day, except for sleep phases, when abrupt transitions between slow-activity sleep and rapid eye movement (REM) occurs. An EEG recording of REM sleep is similar to those with high arousal but reflects episodes of dreaming. Alpha arousal is also called *relaxation.* It is the counter-response to stress and necessary for the body to recharge energy. When a person is in a relaxed physical state, alpha rhythms will be predominant in EEG recordings. Alpha rhythms are used in neurofeedback therapies for children with attention deficit disorder (ADD), whose lack of attention and focus leads to poor learning skills.

3 Arousal and Cognitive Performance

Yerkes and Dodson propose an inverted U-shaped function for modeling the relation between arousal and cognitive performance [15]. The so-called Yerkes–Dodson law consists of two parts (see figure 1). The first part states that cognitive performance increases with arousal up to a task-specific optimum, further decreasing with higher levels of arousal. The second part states that the optimum level of arousal for a given task is inversely related to task difficulty.

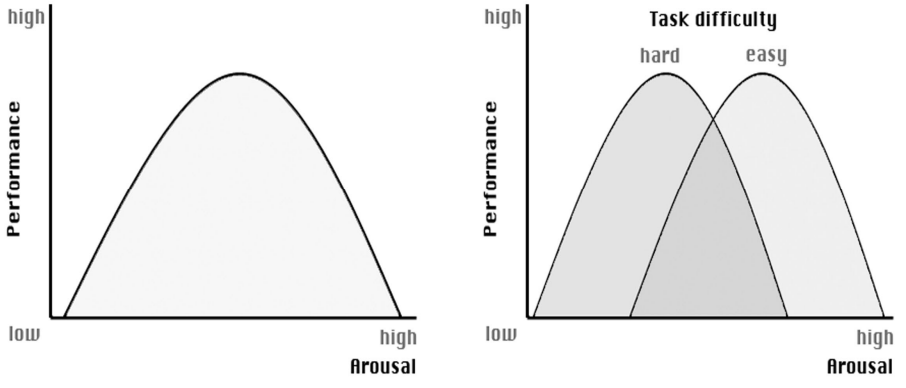


Fig. 1. The Yerkes–Dodson law [15], [16]

Optimal task performance is likely to be achieved in states of high alertness, rather than in states of high arousal. The upward part of the Yerkes–Dodson law can be interpreted as the energizing effect of arousal: increasing arousal improves alertness, which in turn improves performance. The downward part of the Yerkes–Dodson law can be interpreted as the negative effects (e.g. 'tunnel vision') of high arousal (or stress) on task performance. The second part of the Yerkes–Dodson law states that the optimum level of arousal decreases with increasing task difficulty. Cognitively demanding tasks may require a lower level of arousal for optimal performance to facilitate concentration, whereas tasks demanding prolonged efforts or persistence may be performed better with higher levels of arousal in order to increase motivation.

Cognitively demanding tasks can be regarded sensitive to disruption, for example by anxiety induced by aversive stimuli (e.g., a spider) and associated physiological effects, like rapid heart beat or sweating. Arousal and alertness conflict in these scenarios: Arousal can increase cognitive alertness but, as a side effect, cognitive performance suffers from sympathetic nervous system activation [13], [17]. Studies showed positive effects of optimal arousal on learning performance [18], [14].

4 Psychophysiologic Methods and Usability

Psychophysiologic measures are indicators for emotional and attentional responses. A response stimulus can be a startle tone, an emotionally loaded picture, video, or task presented to a subject. Measures of modern psychophysiology include brain activity, skin conductance level (SCL), heart rate (HR), muscle responses, and eye movement measures. Muter et al. [19] demonstrated that psychophysiologic measures are useful for usability studies. They reported about the use of HR and SCL in an experiment with a two-factor design: Each of their 25 subjects was presented a user-friendly and a user-hostile variant of a software tool for accomplishing banking transactions. In order to account for potential differences in task complexity, Muter et al. presented to their subjects memory tasks of varying difficulty in addition to the banking transaction (i.e., problem-solving) tasks. The authors found that the HR was higher during the interaction tasks than during memory tasks (see figure 2). Moreover, they found that HR was hardly related to task difficulty, whereas the SCL was significantly higher during the difficult (i.e. user-hostile) interaction task. Their HR finding can be interpreted as parasympathetic withdrawal, while their SCL finding may suggest that the user-hostile software leads to sympathetic excitation of the sort associated with a fight-flight-freeze response (FFF). Muter et al. conclude that the SCL may be regarded a good indicator for the overall usability of software.

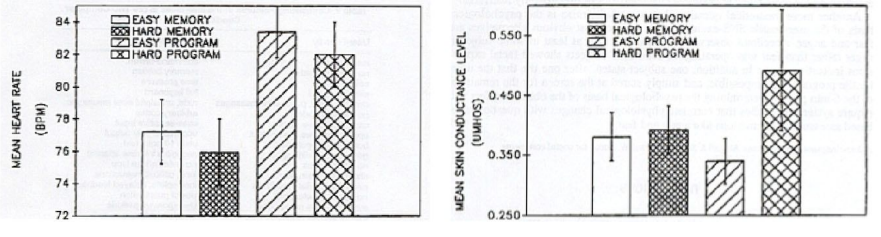


Fig. 2. This experiment from Muter et al. [19] suggests that *SCL (right side)* is a good indicator for usability, as opposed to *HR (left side)*

Both SCL and EEG can show the activation of the sympathetic nervous system. The fight-flight-freeze response (FFF) for example, can be identified in EEG recordings by a dominance of high beta and gamma brain waves. FFF rises cognitive alertness (arousal), as a side effect, however, it reduces the performance of certain higher cognitive functions including articulation, learning, decision making, and interpretation of paralinguistic signals.

From the point of view of evolutionary biology, FFF was key for mastering life-threatening situations. In today's software usability scenarios, however, the activation of FFF is counterproductive.

5 Hypothesis and Experimental Setting

Arousal is the ability to mobilize energy to meet a particular goal. Arousal may be affected by the look and feel of a software interface [20], [21]. Following the Yerkes-Dodson law, the level of arousal is correlated with cognitive performance when using information systems [16]. Specific types of brainwave activity recorded by EEG can be associated with levels of arousal [22], [23], [24]. Against this background, we formulated our working hypothesis: *Changes in the EEG recordings of a subject during a usability test can be used for assessing the learnability of the software used.*

Our experiment followed a two-phase design. In both phases, each subject (N=32) solved a memory game eight times. In the first phase, we used a conventional game with eight randomly presented pairs of pictures, four of them showing abstract and four realistic motives. There were eight different sets of pictures, each of them containing eight pictures. Subjects' task was to find all pairs of pictures. The memory game we used in the second phase used the same pairs of pictures, but the placement of abstract and realistic motives remained unchanged for the eight games in the second phase. We expected that as soon as our subjects discover this non-random placement strategy, they quickly focus their search accordingly, thereby reducing the number of uncovered non-matching pairs of pictures (i.e., error condition). The first phase of our experiment served as control condition. Throughout our experiment, we recorded subjects' brainwave activity with IBVA's EEG system (www.ibva.com). This relatively unobtrusive scanning equipment can be used in laboratory and field settings as well (see figure 3). Due to limitations on the number of electrodes used, IBVA's headband comprises merely 3 electrodes, so this system can record an EEG of the frontal lobe only. For our research, however, this limitation seems dispensable. All signals recorded by the IBVA system are sampled at a 512 Hz rate.

Fig. 3. The IBVA EEG system uses a 3-electrode headband that wirelessly communicates with the IBVA software running on a personal computer

After our experiment, we analyzed the recorded brainwaves according to our working hypothesis as follows: *We consider a software as learnable in case of alpha (8-12 Hz) wave dominance in subjects' EEG patterns. In case of high beta and gamma (25-40 Hz) dominance, however, we consider a software as hardly learnable.*

Following this, we expect that the recorded brainwaves of good learners show a significant alpha dominance, whereas the recordings of poor performers show a high beta/gamma (henceforth called betaH) dominance.

6 Results and Discussion

Following our aforementioned working hypothesis, we aimed at finding a correlation between subjects' individual learning progress and their alpha/betaH brainwave activity. In order to focus on those subjects with the highest and poorest individual learning progress, we selected N=10 subjects for each group, henceforth called 'top players' and 'poor players', respectively. For each game in the two test phases (henceforth called 'control scenario' and 'learning scenario'), we averaged the signal activity in the alpha and betaH frequency bands over all subjects of the particular group. Preceding to this, we averaged the signal activity of both brain hemispheres for each subject. Figure 4 shows the results of our analysis for the control scenario, where pairs of pictures are randomly selected from the card deck at the beginning of each game. On the left chart, we see the alpha signal activity for our learner groups, and on the right chart, the corresponding betaH signal activity is depicted.

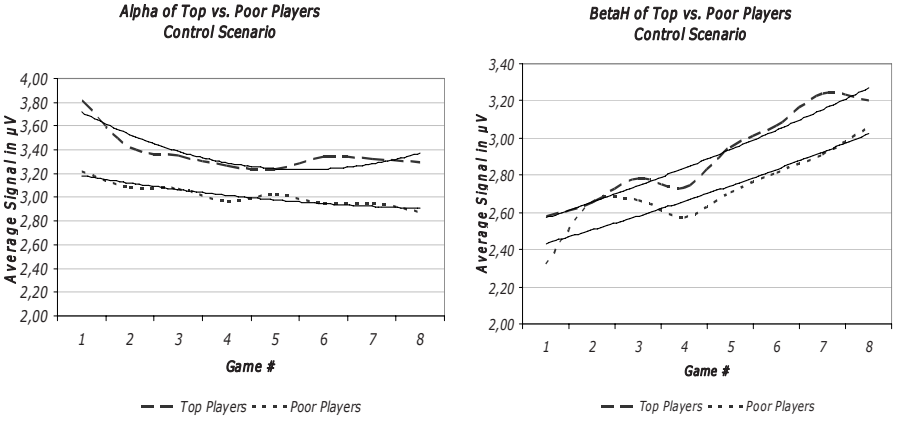


Fig. 4. Alpha and BetaH activity for top vs. poor players in the control scenario

Signal activity in the alpha and betaH frequency bands seems to be quite similar for both groups: apart from a difference in the level of signal activity between top and poor players, both charts show a decreasing (increasing) signal level over the eight games of our control scenario. The polynomial trend lines underline this observation.

The decreasing alpha signal activity over the eight games can be interpreted with the diminishing degree of relaxation for both top and poor learners, albeit on a different signal level.

The increasing betaH signal activity seems to reflect the increasing concentration efforts of both learner groups, again for top players on a higher level of signal activity. Due to our random selection strategy for pictures from the card desk, we

didn't expect to see meaningful learning progress, which, following Yerkes–Dodsons' law, might have reduced the mental efforts (and therefore the state of arousal) necessary for completing the random gaming task at hand.

Figure 5 depicts the results of our analysis for the learning scenario, where the placement of abstract and realistic picture motives remained unchanged for all eight games. Again, alpha signal activity is depicted on the left chart and corresponding betaH signal activity on the right chart.

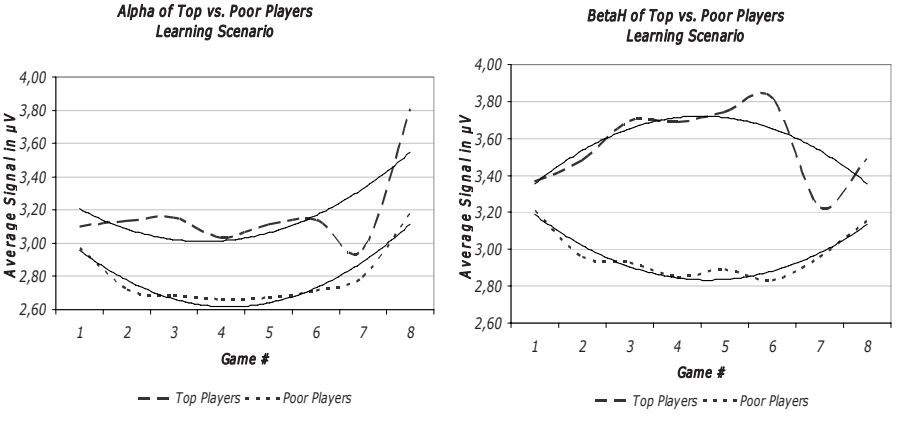


Fig. 5. Alpha and BetaH activity for top vs. poor players in the learning scenario

Now, average signal activity considerably differs from the one we presented so far. In the alpha band, the signal activity over the eight games now seems to form a U-shaped function for both groups, the polynomial trend line emphasizes this interpretation. And again, the signal level is higher for the top players, except for the seventh game, where their signal activity seems to plunge down. The betaH chart shows an even more differentiated recording for top vs. poor players. For the poor players, we see a U-shaped function which is quite similar to the one we have seen for their alpha. For the top players, however, we see an inverted U-shaped function, underlined by the polynomial trend line, with a maximum signal activity during the sixth, followed by a minimum activity during the seventh game.

Having the preceding games of the control scenario in mind, we suspect that the little variation in alpha signal activity up to and including the sixth game can be accounted to the more or less stable status of relaxation in both learner groups. For the last two games, however, we associate the significantly rising alpha with the fact that most users discovered the non-random placement strategy after the sixth game. This assumption is consistent with both users' number of errors when uncovering non-matching pairs of cards and their qualitative feedback during this phase of the game. The significant rise of alpha seems to be different in extent, though, with a steeper increase for the top players.

Regarding betaH, it seems that our poor learners don't show meaningful (or even reduced?) concentration efforts, whereas increasing efforts can be suspected for the top learners up to and including the sixth game (see above). So far, top learners'

investment is quite comparable to their former one during the control scenario. And as already mentioned above for alpha, top players' average betaH signal significantly changes during the last two games; this time inversely, showing reductions in average signal activity.

In order to investigate this a little further, Figure 6 depicts alpha and betaH average activity only for the top players in our learning scenario. Now, the simultaneous decrease for the seventh game, followed by the subsequent increase for the eight game in both frequency bands becomes more clearly visible.

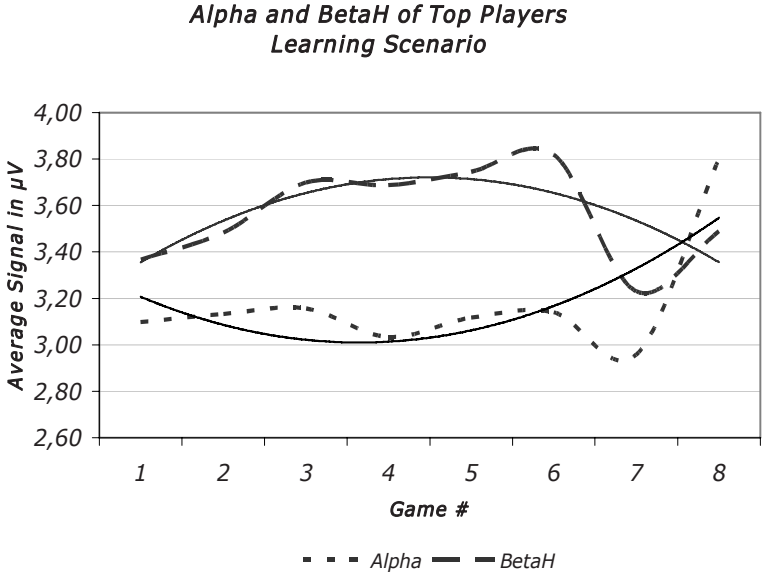


Fig. 6. Alpha and BetaH activity for top players in the learning scenario

The turning point after the sixth game, where all top players reportedly discovered the non-random placement of card pairs may be interpreted as a reduction of task complexity (cf. the Yerkes–Dodson law), which requires less mental efforts (and therefore lower states of arousal) for completing the eased gaming task. The raise of signal activity during a *flash of wit* has been reported by Machleidt et al. (1987). Following this, the subsequent alpha peak may be interpreted as a reward response.

7 Conclusion and Future Outlook

Our experiments showed some interesting, promising and significant differences in brainwave recordings of top performers versus weak performers. The correlations found, however, do not fully comply with our working hypothesis, especially regarding the dominance of alpha brainwaves in the EEG recordings of good learners.

Consequently, we must do further work based on our lessons learned. Limitations within our experimental setting include the (possibly) non-challenging amount of mental efforts necessary for completing our tasks, the small learning effect in the second phase of our experiments, potential noise stemming from a moderately controlled test environment, and the small sample size. Nevertheless, introducing rapid methods of biological usability testing can augment existing usability engineering methods [25] towards rapid biological testing [26]. In order to achieve this, further research is necessary, especially in real-life usability testing and on the systemic level of technology [27]. With the advent of new sensors and input devices, more psychophysiologic indicators of subtle responses will be available for adapting devices, interfaces, and interaction to heterogeneous user needs. Biological rapid usability testing may also offer a new set of methods and tools for software engineers, which can be easily integrated in the development process, at moderate costs in terms of efforts and time. This can further enhance to reach Universal Access in HCI, which aims at "*a conscious and systematic effort in order to advance a proactive approach towards interactive products and environments which should be accessible and usable by the broadest possible end-user population, anytime and from anywhere, without the need for additional adaptations or specialised (re-) design.*" [2], [28]

References

1. Stephanidis, C.: User Interfaces for All: Concepts, Methods and Tools. Lawrence Erlbaum, Mahwah (NJ) (2001)
2. Stephanidis, C., Salvendy, G., Akoumianakis, D., Arnold, A., Bevan, N., Dardailler, D., Emiliani, P.L., Iakovidis, I., Jenkins, P., Karshmer, A., Korn, P., Marcus, A., Murphy, H., Oppermann, C., Stary, C., Tamura, H., Tscheligi, M., Ueda, H., Weber, G., Ziegler, J.: Toward an Information Society for All: HCI challenges and R&D recommendations. International Journal of Human-Computer Interaction 11(1), 1–28 (1999)
3. Squire, L.R.: Encyclopedia of Learning and Memory. Macmillan, New York (1992)
4. Thompson, L.W., Obrist, W.D.: EEG Correlates of Verbal Learning + Overlearning. Electroencephalography and Clinical Neurophysiology 16(4), 332 (1964)
5. Nunez, P.L.: EEG. In: Ramachandran, V.S. (ed.) Encyclopedia of the Human Brain, pp. 169–179. Academic Press, San Diego (2002)
6. Roth, G., Schwegler, H., Stadler, M., Haynes, J.-D.: Die funktionale Rolle des bewußt Erlebten. Gestalt Theory 20, 186–213 (1998)
7. Kolb, B., Whishaw, I.: Introduction to Brain & Behavior. Worth Publishers, Cranbery (2001)
8. Duffy, E.: The psychological significance of the concept of arousal or activation. Psychol. Rev. 64, 265–275 (1957)
9. Marrocco, R.T., Field, B.A., Arousal.: In: Ramachandran, V. S (ed.) Encyclopedia of the human brain, Academic Press (2002)
10. Geen, R.G., Pigg, R.: Interpretation of Arousal and Its Effects on Motivation. Journal of Social Psychology 90, 115–123 (1973)
11. Brehm, J.W., Self, E.A.: The Intensity of Motivation. Annual Review of Psychology 40, 109–131 (1989)

12. Holzinger, A.: Effektivität von Multimedia - Motivation, Aufmerksamkeit und Arousal. GMW FORUM, Zeitschrift der Gesellschaft für Medien in der Wissenschaft,1(00) 10–13 (2000)
13. Holzinger, A.: Basiswissen Multimedia Band 2: Lernen. Kognitive Grundlagen multimedialer Informations Systeme Vogel, Würzburg (2000) http://www.basiswissen-multimedia.at
14. Holzinger, A., Pichler, A., Maurer, H.: Multi Media e-Learning Software TRIANGLE Case-Study: Experimental Results and Lessons Learned. Journal of Universal Science and Technology of Learning 0(0), 61–92 (2006)
15. Yerkes, R.M., Dodson, J.D.: The relation of strength of stimulus to the rapidity of habit formation. Journal of Comparative Neurology and Psychology 18, 459–482 (1908)
16. Holzinger, A.: Multimedia Basics, Learning. Cognitive Fundamentals of multimedial Information Systems. Laxmi, New Delhi vol. 2 (2002) http://www.basiswissen-multimedia.at
17. Brown, V.J., Bowman, E.M.: Alertness. In: Ramachandran, V.S. (ed.) Encyclopedia of the human brain, Academic Press, San Diego (2002)
18. Holzinger, A., Pichler, A., Almer, W., Maurer, H.: TRIANGLE: A Multi-Media test-bed for examining incidental learning, motivation and the Tamagotchi-Effect within a Game-Show like Computer Based Learning ModuleEducational Multimedia, Hypermedia and Telecommunication 2001. Association for the Advancement of Computing in Education, Charlottesville (VA) pp.766–771 (2001)
19. Muter, P., Furedy, J.J., Vincent, A., Pelcowitz, T.: User-Hostile Systems and Patterns of Psychophysiological Activity. Computers in Human Behavior 9(1), 105–111 (1993)
20. Norman, D.A.: Cognitive engineering. In: Norman, D., Draper, S. (eds.) User Centered System Design: New Perspectives on Human-Computer interaction, Lawrence Erlbaum Associates, Mahwah, NJ (1986)
21. Norman, D.A.: Emotional Design: Why we love or hate everyday things. Basic Books, New York (2004)
22. Gutjahr, L., Machleidt, W.: Influences of Psychic Variables on the EEG During a Conflict-Centered Interview. Electroencephalography and Clinical Neurophysiology 48(3), 25–26 (1980)
23. Hinrichs, H., Machleidt, W.: Basic Emotions Reflected in EEG-Coherences. International Journal of Psychophysiology 13(3), 225–232 (1992)
24. Bruggenwerth, G., Gutjahr, L., Kulka, T., Machleidt, W.: Music Induced Emotional EEG Reactions. Eeg-Emg-Zeitschrift Fur Elektroenzephalographie Elektromyographie Und Verwandte Gebiete, 25(2), 117–125 (1994)
25. Holzinger, A.: Usability Engineering for Software Developers. Communications of the ACM 48(1), 71–74 (2005)
26. Holzinger, A., Slany, W.: XP + UE -> XU Praktische Erfahrungen mit eXtreme Usability. Informatik Spektrum 29(2), 91–97 (2006)
27. Holzinger, A., Leitner, H.: Lessons from Real-Life Usability Engineering in Hospital: From Software Usability to Total Workplace Usability, In: Holzinger, A., Weidmann, K.-H (eds.) Empowering Software Quality: How can Usability Engineering reach these goals? Austrian Computer Society, pp. 153–160 (2005)
28. Stephanidis, C., Savidis, A.: Universal Access in the Information Society: Methods, Tools and Interaction Technologies. Universal Access in the Information Society 1(1), 40–55 (2001)

Learning Through Exploration, Autonomy, Collaboration, and Simulation: The 'all-in-one' Virtual School of the Hellas Alive!© Online, Language-Learning Platform

Dimitris Tolias[1] and George Exadaktylos[2]

[1] Hellenic American Union
22, Massalias St.,
10680 Athens, Greece
dtolias@hau.gr
[2] Hellenic American Union
22, Massalias St.,
10680 Athens, Greece
gexadaktilos@hau.gr

Abstract. The Hellenic American Union (HAU), in Greece, in collaboration with the Academy of Electronic Media at the Rensselaer Polytechnic Institute, in the USA, has developed Hellas Alive!©, an online language and culture learning platform for teaching Greek as a Foreign Language with emphasis on interactivity for language learning, cultural awareness, and communication. The platform comprises the Virtual School component, the Virtual City component and the cultural Playspaces component. Language learning takes place in the Virtual School which supports autonomous exploration, simulation, representation and collaboration, with rich multimedia content and synchronous and asynchronous learning tools. Its interface includes an expandable playlist that provides users with all the tools necessary for effective language learning. Such tools include unlimited options to read text, to listen to audio, to watch video or to see images, to record, to submit audio and text, to perform 23 different types of exercises with instant automated error feedback, to chat with other users, and to share material with other users. The platform also includes a variety of interfaces with additional information on language and vocabulary. Users can activate them every time they encounter the corresponding grammatical item or word.

Keywords: Modern Greek, language learning, exploration, autonomy, collaboration, simulation, online.

1 Introduction

The Hellenic American Union (Greece) and the Academy of Electronic Media of the Rensselaer Polytechnic Institute (U.S.A.) have been cooperating, since 2000, on a research project concerning the design, development, and completion of the online

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 823–832, 2007.
© Springer-Verlag Berlin Heidelberg 2007

platform Hellas Alive!© for the distance learning of the Modern Greek language. Hellas Alive!© comprises the following areas: the Virtual School, with access to lexical and grammatical information, the 3D Virtual City[1] and the interactive game-like cultural Playspaces[2]. During its present phase of development, the platform aims at the long-distance teaching of Modern Greek as a Foreign Language to learners of beginner to lower-advanced levels[3], with a strong emphasis on interactive communication and cultural information.

2 The Hellas Alive!© Virtual School: Overall Description

The Hellas Alive!© Virtual School constitutes the main learning area of the application. Learners in the Virtual School access the learning material through a single interface. All language skills[4] are practiced within this interface and a number of activities are supported through the same environment. Learners are able to perform the following tasks: *authentic text reading and listening with hyperlinked information; image and video viewing; selection among 23 different exercises and activities in expandable text; own speech recording; performance evaluation; audio submission; tutor contact through the in-built chat and contact tutor features; grammar and syntax database access with tips on appropriate contextualized language usage and automated feedback on correct answers*.

3 Learning Through Exploration

On the one hand, according to Kellogg et al [9], "…user interfaces that support learning by exploration are well-known, successful methodologies…" On the other, according to Bhavnani and John (in Kellogg et al [9]) "the agenda for HCI needs to

[1] The 3D Virtual City is the main area for cultural immersion and can be accessed directly from the Virtual School interface. It is a virtual interactive environment that reenacts a contemporary Greek cityscape, using authentic colorful images. Equipped with fully functional navigation tools, the Virtual City provides users the option to choose 'where to go' and – through rollovers, associated with embedded audio files – to hear in Modern Greek names of stores and buildings. If users want to become more familiar with Greek culture and civilization they have access to the interactive cultural Playspaces, through the Virtual City interface. Also, its toolbar gives users access to in-built communication tools that allow them to communicate both with each other (synchronously, through a 'chat' feature) and their tutor (asynchronously, through a 'bulletin board' feature).

[2] The interactive cultural Playspaces aim at raising users' awareness on Greek culture and civilization and at introducing them to related language, with a variety of degrees of interactivity. To the present, the following eight generic Playspace types have been created: "How To", "Art Exploration, "Documentary", "Maps and Directions", "Meet the Parea", "Word Jumble", "Let's Talk Greek", "Scavenger Hunt".

[3] According to the Common European Framework of Reference for Languages of the Council of Europe (Council of Europe [5]), the present phase of Hellas Alive!© aims at teaching Modern Greek to learners of A1 to C1 levels of competency.

[4] Reading, writing, speaking, listening.

shift away from support for learning by exploration to support for training users on productive strategies."

No matter whether one claim is more valid than the other, according to White et al [19], when people perform an exploratory search[5], typically, they are uncertain about the area within which they are searching (the 'where' aspect of the exploration) and about the reason that triggered their search (the 'why' aspect of the exploration). Also, they "engage in exploratory searching both intentionally and incidentally to other activities." Such exploration involves situations in which people:

a. do not know at all or know partially what they are searching for

b. may know what they are searching for at the beginning of their search but, later, when they are confronted with new information, they lose this knowledge, and

c. "recognize useful information objects by scanning through information resources, evaluating their usefulness and determining the content/structure of a set of information objects."

Situations of, mostly, the 'a' and 'b' types (but, to a somewhat lesser degree, of the 'c' type as well) suggest that interface design is crucial to the facilitation of user exploration activities. For example, as Shneiderman [15] supports, one of the characteristics of a successfully designed interface should be the option to easily reverse actions, since it would make users less stressed and allow them to explore unfamiliar options. To this, it can be added that automated feedback on user actions would also make users less hesitant to explore.

Furthermore, it can be argued that interfaces with high degrees of interactivity encourage exploration through option to access a variety of tools and functionalities, through immediate response to user actions, through active learning (in the case of e-learning systems or other learning programs), and through users' involvement with the content. Needless to say, when interfaces provide unified access to information, exploratory users' search activities are facilitated.

3.1 Exploration Through Unified Access in the Hellas Alive!© Virtual School

According to Zeldman [20], "the so-called 'Three-Click Rule,'…states that users should ideally be able to reach their intended destination within three mouse clicks". Zeldman agrees with the rule in principle, but his view is not entirely adopted by Porter [13]. His research shows that "the Three-Clicks Rule is just a myth" and that "the satisfaction of users doesn't depend on the number of clicks". According to his findings, what matters is not perhaps the number of clicks that users make in order to reach their desired destination but mostly the degree to which they are successful in reaching it. To this, Zeldman [20] suggests that designers should create interfaces with a logical structure and adhere to the "Rule of Five [that] postulates that complex, multi-layered menus offering more than five main choices" may confuse users.

No mater which or whose point of view is correct, in the Virtual School, the number of clicks that are necessary for learners to reach any type of information is

[5] Even though White et al [19] refer to searches performed through search engines, bibliographic databases and digital libraries, their claims can apply easily to searches performed within interfaces that encourage exploration, such as the ones of the Hellas Alive!© platform, and, in particular, the Virtual School area.

never more than three; each time earners wish to do a certain exercise or perform a certain activity, they do not have to switch interface. Even in the case that learners want to access a grammar or vocabulary entry it will never be more than one click they will make. As for layers, there are only two, in the Virtual School: the one of the main interface with the toolbars, media screen, and playlist, instructions, and workbook areas and the ones of the grammar and vocabulary entries interfaces.

4 Learning Through Autonomy

It could be argued that the advent of the personal computer and the Web in the early 90s revolutionized language teaching (Hirvela [7]) and accelerated the efforts and experimentation in distance education. As Kern [10] notes, computer networks, both local and worldwide, provide possibilities for new interpersonal contacts, and communicative engagement among students supporting their efforts to explore the foreign language.

Computer communication in the educational context is founded on the principle of bringing students together while enhancing their autonomy and independence. Fey [6] observes that computer enhanced learning environments allow students to "transcend boundaries of classroom walls and to learn in new ways". Learners are thus no longer confined to the traditional types of classroom based activities and interaction with their peers that they may find boring, and counterproductive in the 21st century. The creation of a 'distancing effect' supported by electronic communication seems to have lent itself to increased student empowerment and desire to express themselves in the new language since the computer seems to minimize the risks involved in the communication and usage of a foreign language.

Over the last two decades, the concepts of learner autonomy and independence have gained momentum. Wenden [18] claims that autonomy has been the premium placed on the role of the learner in the language process. This shift in the responsibility from teachers to learners does not exist in vitro, but instead it is the result of a concatenation of changes to the curricula leading towards a more learner-centered type of learning. What is more, this reshaping of teacher-learner roles has been conducive to a radical change in the age-old distribution of power and authority we are all so familiar with in the traditional classroom learning environment. Under the perspective of autonomy, learners are expected to assume greater responsibility by taking control of their own learning. This means that learner autonomy is a dynamic process rather than a mere static representation of a learning environment. Technology helps learners assume this power and gain control over their own learning, allowing them to become aware of and identifying new learning strategies.

Little [11] defines autonomy as essentially a matter of the learners' psychological relation to the process and content of learning, a capacity "for detachment, critical reflection, decision-making, and independent action". One could also argue that the independent learner assumes a proactive role in the learning process generates ideas and avails himself of learning opportunities, rather than simply reacting to the stimuli and feedback of the teacher. Within this framework, learning is not simply a matter of mechanical memorization but a "constructive process which involves actively seeking meaning from events (Candy [2]).

Learner autonomy has been defined as learners taking responsibility for their own learning (Usuki [17]). Learner autonomy in other words, seems also to denote learners' internal attitude towards themselves as learners. Furthermore, autonomy does not mean giving over the learning environment to learners, but instead it requires that learners are motivated to create their own environment for learning. In other words, the learners' acceptance of responsibility seems to be the basis of learner autonomy, which in its turn has both socio-affective and cognitive implications that need to be taken into consideration.

Technology is considered a major tool that provides learning autonomy (Sanz et al [14]). This necessitates a specific formation to emphasize that it is important to 'learn how to learn'. To begin with, physical autonomy seems to be an area to which computers contribute most as far as the learning process is concerned with the strong assumption that computers are beneficial for autonomous learning. Also, computers permit learners to study parallel to their rhythms, independent from others all being active of solicitations of the machine. Moreover, because of the re-dramatization of error via constructive feedback, and modification of the relationship of teacher-learner interaction, a great degree of motivation may often be observed. Technology offers also consultation typologies in the context of learning (Cord-Maunory [4]). The computer, for instance, supports the feature of self-evaluation in the form of direct or indirect feedback sessions. Moreover, it supports guidance, meaning the unity of means implemented in order to advise, orientate, inform and lead learners at the time of their interactions with the computer.

It could be argued that this guidance can be divided into two fundamental formats: the functional guidance and the pedagogical guidance. The former refers to the unity of components dependent on the functionality of buttons, colors etc, which help to orientate learners in the course. The latter is the unity of components provided by the designers (material developers and software developers alike) which guide learners and lead their reflections over the content. These two types of guidance are usually provided by a series of orders and instructions own to each exercise. These in turn are usually clear and short pointing out what learners have to do each time (e.g. fill in the blanks, drag and drop, etc). Interactivity, on the other hand, seems to be executed by the navigation and the application of instruction and orders such as "select the correct answer" or "choose the right answer". Further to that, there are also the complementary aids which are defined as background components meant to help learners in their autonomous study. These present themselves outside of the context conforming to the particularities of each learning application. Therefore, a successful learning platform that uses technology to accelerate autonomous learning needs to combine all the above aids. Lack of structure and clear indications over major pedagogical presuppositions such as objectives, level and target audience may disorient learners in their efforts to master the foreign language.

4.1 Autonomy in the Hellas Alive!© Virtual School

In contrast to the traditional classroom teaching environment, learning a foreign language in the Hellas Alive!© Virtual School is achieved without the need for an instructor to be physically present. In other words, the role of the instructor is redefined, since care has been taken for the platform itself to be aided and not guided

by an instructor's presence, throughout the learning process. However, as learner's need for contact with the human element is taken for granted, learners are given the opportunity to contact a tutor through the in-built contact tutor feature. Also, they are given access to the learning material exactly as they would, if they were taught in a classroom environment, but, if they choose, they can follow their own learning pace and learning style.

5 Learning Through Simulation

Simulated learning in a computer enhanced environment is considered a promising tool for supporting learning at all levels (Kartal [8]). Learning through simulations, either by creating simulations or by exploring existing ones helps learners develop a deeper and better understanding of the concepts being simulated. For instance, exploring simulations created by others exposes learners to ideas that may be different from their own, and gives them the opportunity to experiment with those ideas. A simulation is also a tool to grasp ideas and concepts which are already known to the teacher but are difficult to understand for the average student. Learners seem to follow a standard process when working with simulations; they state the hypothesis, they devise and run simulation experiments and analyze the results. Moreover, they interact with ideas when they view and attempt to explain simulation results. Thus, working with simulations can promote collaboration among learners because simulations are concrete and can be manipulated and evaluated.

The results of a simulation may also yield different interpretations which can then be discussed within a group of learners. This exchange of ideas among group members can enhance the meaning that learners derive from the simulation activities. Interactive simulations may also accommodate a range of engagement for learners, depending on the level of interest and of the time available. The educational benefits of this situation are probably similar to watching a video in class or listening to an audio text. In a more time-consuming mode, learners may also experiment in pro-created simulations. This type of activity allows learners to explore a topic by comparing or contrasting different scenarios participating thus interactively into the action.

5.1 Simulation in the Hellas Alive!© Virtual School

The fundamental and innovative nature of the Hellas Alive!© learning platform is the integration of a broad and highly varied range of web-delivered, feedback rich, interactive language-learning and cultural awareness components within a highly stimulating learning environment. Following the simulated learning philosophy, learners are introduced to real world contexts and interactions with the use of authentic texts and a variety of exercises and activities that are presented and executed on line. Learners are also immersed in a contemporary Greek and Greek-American cultural context enhanced by a variety of media such as pictures, sound files and videos.

6 Learning Through Collaboration

According to Chickering et al [3] the ideal learning environment "…encourages student-tutor contact, promotes cooperation among students, supports active learning, provides prompt feedback, emphasizes time on task, communicates high expectations and respects diverse talents and ways of learning…"

Collaborative theories of learning view learning as a socially interactive process that is enriched by interpersonal communication. This construct emphasizes the importance of knowledge that students bring to their new experiences and values of shared perspectives which, in effect, draw substance from different ways of knowing (Mezirow [12]). Learning is not a result of development; learning is development. It requires invention and self organization by the learners. Learning a language may require the undoing of earlier conceptions. Therefore, disequilibrium facilitates learning. To this effect, learners in the Hellas Alive!© virtual communities benefit from challenging, open-ended investigations in realistic and meaningful contexts seeking to organize and synthesize their experiences. Consequently, dialogue within a community engenders further learning. Ideas are accepted as truth only when they make sense to the community and thus rise to the level of shared understanding.

Technology helps to this direction since it is the most expedient vehicle for change. Technology embraces distance learning to allow flexibility in meeting learners' time constraints and not only. Time is viewed as a resource and emphasis is placed on learners' performance and activities. Learners have thus the opportunity to learn from and teach one another (Angelo [1]).

6.1 Collaboration in the Hellas Alive!© Virtual School

Hellas Alive!© helps build communities that enable teamwork and interpersonal involvement. The on line community that Hellas Alive!© develops, promotes a culture of inquiry and evidence, with explicit, broadly shared goals for learning outcomes. This is a culture that encourages cooperative and collaborative methods for individual advancement, resulting in a more effective and reinforcing community. Through participation and involvement learners apply relevant knowledge to improve practice.

7 Integration in the Hellas Alive!© Virtual School

In the Virtual School learning interface, there are multiple screens serving the same single purpose, which is to educate users. For example, the Virtual School interface provides an expandable text screen (with "playlist" options), three text screens (one provides automated bilingual instructions, the other links to related grammar and vocabulary, and the third text of dialogues presented in respective lesson), and one video screen (with associated images or video). Also, there is consistent linkage between all the platform's components and the Virtual School's individual parts. For example, in the screen with the presented dialogues text, potentially unknown vocabulary or grammatical items are linked with their respective areas. Last, toolbar buttons provide vertical and horizontal access to the other platform interfaces. HCI

theory supports such integration, when applied for educational purposes. In the case of Hellas Alive!©, sensing, rendering, and simulation (Stappers et al. [16]) are integrated and create the impression of regular classroom learning situation.

8 Conclusion

As both theory and practice demonstrate, an interactive, language e-learning tool that encourages learners' exploration, collaboration and autonomy and simulates natural-learning situations is expected to have positive learning outcomes. This seems to be the case of the innovative Hellas Alive!© platform as well, which, upon completion of all materials development, will comprise five levels of Modern Greek competency.

Currently, material for the first level has been completed and piloted, and material for the second level is soon to be completed as well. The platform is expected to be available to the public during 2007.

References

1. Angelo, T.A.: The Campus as Learning Community. American Association for Higher Education 49(9), 3–6 (1997)
2. Candy, G.: Self Direction for Lifelong Learning. Jossey-Bass, California (1991)
3. Chickering, A.W., Gamson, Z.F.: Seven Principles for Good practice in Undergraduate Education. In: The Wingspread Journal, vol. 9(2) (1987) Special Section
4. Cord-Maunoury, B.: Analyse du Site Lire-francais. In: ALSIC, vol.2(4), pp. 239–254 (December 1999)
5. Council of Europe. Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Retrieved November 20, 2006 from the World Wide Web: http://www.coe.int/t/dg4/linguistic/CADRE_EN.asp
6. Fey, M.: Critical Literacy in School-college Collaboration through Computer Networking: A Feminist Research Project. Journal of literacy Research 30(I), 85–117 (1998)
7. Hirvela, A.: Computer-mediated Communication in ESL Teacher Education. In: ELT Journal 60(3) (2006)
8. Kartal, E.: The Internet and Autonomous Language Learning: A Typology of Suggested Aids. University of Uludag (2005)
9. Kellogg, W.A., Clayton, L., Polson, P.: Introduction to this Special Issue on New Agendas for Human-Computer Interaction. In: Human-Computer Interaction,15(2&3) (2000)
10. Kern, R.G.: Restructuring Classroom Interaction with Networked Computers: Effects on Quantity and Characteristics of Language Production. The. Modern Language Journal 79(4), 457–476 (1995)
11. Little, D.: Learning as Dialogue: The Dependence of Learner Autonomy on Teacher Autonomy System 23(2), 175–181 (1995)
12. Mezirow, J.: Transformative Dimensions of Adult Learning. Jossey-Bass, San Fransisco (1991)
13. Porter, J.: Testing the Three-click Rule. In: User Interface Engineering (2003) Retrieved November 17, 2006 from the World Wide Web http://www.uie.com/articles/three_click_rule/

14. Sanz, M., Villanueva, M.L.: Le Projet Small et les Enjeux de l'Hypertexte pour l'Autoformation. In : Le Francais Dans Le Monde, Paris : Cle International, Janvier, pp. 66–75 (2002)
15. Shneiderman, B.: Designing the User Interface, 3rd edn. Addison-Wesley, Reading (1998)
16. Stappers, P.J., Overbeeke, K., Gaver, W.: Beyond the Limits of Real-time Realism: Moving from Stimulation Correspondence to Information Correspondence. In: Hettinger, L.J., Haas, M.W. (eds.) Virtual and Adaptive Environments, pp. 91–110. Lawrence Erlbaum Associates, Inc, New Jersey (2003)
17. Usuki, M.: Learner Autonomy: Learning from the Student's Voice. In: CLCS Occasional Paper, vol. 60 (2002)
18. Wenden, A.: Learner Strategies for Learner Autonomy, UK. Prentice Hall, Englewood Cliffs (1998)
19. White, R.W., Kules, B., Bederson, B.: Exploratory Search Interfaces: Categorization, Clustering and Beyond. In: Report on the XSI, Workshop at the Human-Computer Interaction Laboratory. University of Maryland (2005)
20. Zeldman, J.: Taking Your Talent to the Web: Making the Transition from Graphic Design to Web Design. New Riders (2001)

Appendix: The Virtual School

Appendix: The Virtual School

1: Playlist
2: Instructions (Greek-English)
3: Images-Videos
4: Video transcripts – Exercises with links to other related grammatical or vocabulary items
5: Links to other related grammatical or vocabulary items
6: Main toolbar with access to: Virtual City, grammar and vocabulary indexes, built-in applications
7: Toolbar with information on unit and lesson numbers and access to 'delete' and 'navigation' functions

Co-learn: Collaborative Learning Engine a Vision for Pervasive Collaboration in E-Learning

Vaibhav Tyagi and Sheikh Iqbal Ahamed

Department of MSCS, Marquette University, Milwaukee, Wisconsin, USA
{Vaibhav.Tyagi,Sheikh.Ahamed}@marquette.edu

Abstract. The quest for knowledge has been an intricate part of the human psyche and, more often than not, we learn by sharing our experiences and combining our efforts. Collaboration is an important part of the learning experience, and when combined with the technological boom of present times, can be a platform for the future of continuous education. In this paper, we present our vision of a collaborative learning engine, which allows the users to collaborate in real time using their mobile devices, such as smart phones and/or PDAs. Our proposed collaborative learning engine will allow handheld devices to communicate among each other and share text, files, audio and video from different sources. In our initial design and development, we have addressed up to eight devices for streaming video and audio content, however other channels of collaboration can be extended for more users.

Keywords: Co-Learn, Collaborated, Collaborative, Pervasive, Learning, Life Long Learning, and Mobile Collaboration.

1 Introduction

In today's highly competitive world, education plays an important role. According to the U.S. Department of Education's [2] projection, 15,105,000 new students will enroll in college in the year 2006. Apart from a college based education, today more than ever there is a need for life long learning [7, 8]. Increased interest in higher education combined with progressive and continuous advancements in technology has paved the way for electronic computer based learning in recent years. A few years ago, laptops were reserved for business travelers, but thanks to the technological race to make devices faster, smaller and cheaper, the trends today have changed. Many universities and even some high schools today require their students to be equipped with laptops or handheld devices (Windows Mobile / Palm / other similar PDAs).

Though traditional teaching techniques have not changed much, the use of technology in a classroom setting is no longer reserved for technical fields. Education can no longer be bound merely by plain printed text. Today we combine and use multiple mediums of information together, such as sound, graphics, and simulation along with traditional media. It is imperative that these mediums work effortlessly in harmony. This for the most part has already been achieved on desktop based computing systems. However, the fast growing handheld market has yet to see such systems. This idea has sparked the interest of some researchers [1, 7] and has lead to

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 833–842, 2007.
© Springer-Verlag Berlin Heidelberg 2007

development of some promising systems, such as X-Media Learning Objects (XLOs). Models of collaboration and use of scripts [11] have also been researched; however, our research differs from their area of focus.

With the availability of broadband internet and high speed networks, students today can access information stored at a remote site and collaborate with each other during the research and learning process. Although information sharing is not new, most handheld- based devices and applications today [9] focus on the actual process of information sharing, rather than its users [1]. We understand it is essential to tend to the needs of the nascent world of handheld devices. But as Holzinger [6] points out, e-learning demands User Centered Design (UCD) techniques instead of System Centered Design techniques. We will discuss some UCD techniques which have already been designed and implemented for handheld devices, and present our vision of a collaborative learning engine called *Co-Learn*. In this paper, we extend the already established importance of user-centric design, by presenting a tool which allows for seamless integration of technology with collaborative learning and teaching.

In section 2 we present the motivation behind this research. Section 3 presents the primary characteristics of collaborative learning. Section 4 discusses the present work in the field. Section 5 describes *Co-Learn* and our approach, and presents an example where *Co-Learn* can be used. We discuss the research challenges for *Co-Learn* in Section 6, followed by conclusions and future work.

2 Motivation

Our motivation for *Co-Learn* comes from varying sources, including our personal experience with the first windows CE based handheld device. When wireless internet on Pocket PCs was still in early development, we had a chance to experiment with it, and found it extremely difficult to setup. Although this is no longer the case with wireless internet, a few other services provided by such devices can be extremely daunting to setup and use. More often than not such services go unnoticed since they are too system centric for even a person experienced in technology to configure and put into practice. Let's look at a few scenarios where the use of user-centric design combined with an integrated solution could be helpful.

Students are often required to collaborate with each other both inside and outside the classroom. Let's take a look at a typical class room example; say a group of four students (Ryan, Dan, Jake and Mike) are working on a software project. Dan and Mike are responsible for gathering requirements, planning the overall design of the software, and preparing the documentation while Jake and Ryan are responsible for implementation. Jake and Ryan both have full time day jobs, while Dan has a long commute to school. Mike works on campus and has prior commitments most evenings. The team can only meet for limited times since their schedules often conflict. Although this example may sound extreme, situations like this are common. The rising costs of quality education has forced students to juggle both work and school, and student expectations have risen as a result of the highly competitive environment. The team is left with few options for project planning and discussion. In attempting to discuss the project, the team decides to meet 10 minutes before and after

the class to delegate responsibilities. If discussion is required, they meet with individual teammates whenever possible and keep other team members informed via email or an instant messaging system.

This proposed plan requires individual team members to keep track of who is doing what, and share the information with others via email or IM. Teammates may not have the most up to date information, or may not understand their responsibilities, since there is no way of directly monitoring the other members' progress. The team can only hope that each of the other team members can deliver what is expected of them when the deadline arrives. Chaos and confusion can be easily avoided with the help of a tool which allows the users to perform their tasks and duties from any location. Such a tool would also allow the users to track the work of their colleagues, share information, assist and provide feedback on work done by others and communicate effectively.

Based on our observations, effective collaboration has the following general characteristics:

C1) *Sharing of resources:* As Yau and his team [9] point out, sharing of knowledge is an important aspect of collaboration and essential for a truly enriching collaborative learning experience.

C2) *Security and access control:* Although we want users to share ideas and concepts, we would like to restrict access to certain files and/or other resources if need be.

C3) *Clearly defined goals/tasks:* Clearly defined goals and tasks form the foundation of organized learning and allow for the team to track its progress, which brings us to C4).

C4) *Tracking progress:* Since different people bring different skills to the table, tracking progress in both quality and quantity may be required.

C5) *Seamless communication:* Since handheld devices have small screen sizes and even smaller keyboards, a more natural form of communication may be more appropriate than text based communication.

C6) *Providing Feedback:* When goals are clearly defined (as in *C3*), proper feedback allows the learner to backtrack to do more research or offer or clarify points.

C7) *Enable Central Organization:* Since modern learning techniques combine rich media with traditional teaching methods, it is essential that such content be organized and be easily accessible by participating learners.

3 Related Works

Collaborated learning is not a new concept. It has been one of the basic foundations of learning and is almost as old as the concept of learning itself. Much research has been done to establish the importance of collaborative learning and how it helps develop analytical skills [5]. Modern society has been driven forward by continuous developments in education and advancements in technology. Researchers have long understood this fact and have been developing various ways to help enhance collaborative learning using pervasive computing.

Holzinger and his/her team have shown the importance of life long learning and how X-MEDIA objects [7] can be used to assist the continuous growth. They talk

about the importance of continuous education and how desktop and laptop computing are not enough to fulfill the "whatever and wherever a learner wants to" [7] concept. They then propose the 'iXmedia Learning Engine' which can be used "for every type of learning technology and learning methods" [7]. This is where Holzinger addresses the sharing of resources (C1). The content is handled by the central content management system which ensures security & access control (C2) and organization (C7). This, however, makes the iXmedia an infrastructural solution, where clients must connect to a dedicated content management system.

In the "Smart Classroom" [9], Yau et al. have presented another approach towards pervasive and collaborative learning, which combines a fixed infrastructure with mobile and pervasive ones. The Smart Classroom provides situation awareness to the mobile devices by using fixed infrastructural components such as sensors. The mobile devices utilize the ad-hoc network to communicate with each other. This is where Yau's team addresses seamless communication (C5) although not to its fullest extent. An ad-hoc network is used to synchronize notes between the students/instructors and/or student's personal computers; this addresses the sharing (C1) and feedback (C6) aspects of collaborative learning. The collaboration model of the Smart Classroom emphasizes sharing educational material such as lecture notes, design documents and student's questions/queries.

Davis proposes a prototype called Mobile Media Metadata for Media Sharing (MMM2) [1]. MMM2 attempts to facilitate the photo sharing experience between handheld devices and cellular phones with embedded cameras. MMM2 automates "*how* a photograph is shared" [1] and provides assistance with other related computation tasks. Although Davis's research does not focus on collaborative learning, similar concepts can be used to share resources C1) during collaborative education.

Collaboration itself is a complex paradigm and can present various challenges. Many researchers have addressed human factors that come into play during the process of collaboration [10] [11]. Their research, however, focuses on topics such as models of collaboration, trust, and conflict resolution based on rules and use of scripts. Although interesting, these topics are out of our present scope of research.

4 *Co-Learn*: Collaborative Learning Engine

As we have already seen, collaborative learning allows students/participants to perform their assigned tasks and effectively communicate with other team members. Benefits of collaborated education have been established [7]. *Co-Learn,* our current project, is a vision of a collaborative learning engine which focuses on the users' needs, and provides them powerful tools with a simple user interface. *Co-Learn* aims to provide the following features to its users:

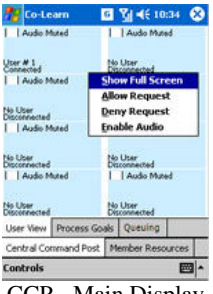
- **Central Command Post:** The Central Command Post (CCP) is primarily the team leader's view of the members; it allows the team leader to monitor the progress of, and co-ordinate with, all other members.
    - The CCP can be used to view up to eight member screens at the same time. This allows the team leader to receive feeds from all members and track team progress (C4).

- The CCP allows the team leader to have eight way video-conferencing, which allows for seamless communication (C5).
- The CCP allows the team leader to zoom into any one of the member's camera feeds. This also extends the concept of progress tracking (C4) and seamless communication (C5).
- The CCP allows the team leader to queue and arrange the audio visual feeds received from different members. This covers portions of the centralized organization requirement (C7); other features cover the rest of this requirement.
- The CCP allows the team leader to send and share files and/or text to individual/multiple members. This addresses s*haring of resources* (C1).
- The CCP allows the team leader to make a list of team goals and post them on the Member Resource Center wall (C3). Each of the items on this list can transition to one of the following statuses.
  - Open – unassigned: This status signifies that the goal has not been assigned to any team member and has not been started.
  - Open – assigned: This status signifies the goal has been assigned to a team member but has not been started.
  - In Progress: This status signifies that the assigned team member has started working towards the goal.
  - Complete: This status signifies that the assigned team member has completed his/her goal to their satisfaction.
  - Needs more work: This status signifies that the initially submitted work has been deemed insufficient and more work is required. This allows the members to provide feedback to each other (C6), and in turn improve the overall quality of work.
  - Closed: This status is restricted for team leader's use only, and signifies the related goal has been completed to the team leader's satisfaction.
- The CCP allows the team-lead to perform various administrative duties, which may include
  - Updating the team goals and their statuses.
  - Interrupting or switching the audio visual feeds.
  - Reassigning roles and duties to different members.
  - Approving requests.

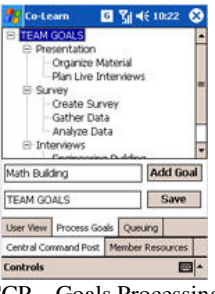
  This allows the team leader to organize the content and manage the workflow, enabling central organization of tasks (C7).

- **Member Resource Center:** The Member Resource Center (MRC) is geared towards non–administrative functions and is focused on team activities. The MRC acts as a central place where team members can get the list of team goals, assign goals to themselves, share documents, and post comments. A document or audio visual feed can be attached with each of the goals when the goal transitions from the 'In Progress' state to 'Complete' state. Some other features of MRC are
  - It allows the members to access the unshared documents of others, while providing security (C2).
  - It allows the members to change permission settings on documents and control the access of their content (C2).

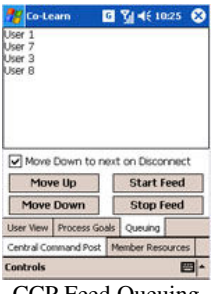
- It informs the members if a new document is posted on the wall, thereby facilitating seamless communication (C5) and central organization (C7).
- It allows the members to communicate (C5) with other members using Voice over Internet Protocol (VoIP).
- It allows the users to request audio visual feed from other members. This feature also addresses access control (C2).
- It allows the team members to view the presentation channel. The users may use this channel either to share the content with others, fulfilling (C1) or track progress (C4) of other members while information over audio/video feed is relayed to the stage view.
- If the channel is occupied, the MRC will also allow the users to queue their feeds to the presentation channel. Team members can request access from the team leader if the channel is unoccupied. This again demonstrates that *Co-Learn* provides clear communication (C5) between collaborating members.
- It allows the team members to update the status of individual goals set by the team, from 'Open' to 'In Progress' and from 'In Progress' to 'Complete', allowing them to track the progress of their goals (C4).
- It allows the team members to view the work of other team members once it is in 'Complete' status. Once a Goal/Task is set to 'Complete' all related documents are automatically shared (C1). If the member reviewing documents feels the need for more work, they can post comments and change the status to 'Needs More Work'. This allows members to track each other's work (C4) and provide feedback to other members (C6).



CCP - Main Display



CCP – Goals Processing



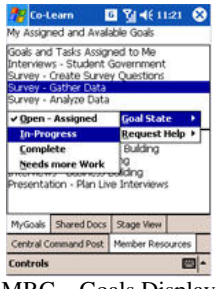
CCP Feed Queuing

Fig. 1. Central Command Post – Screen Shots

Co-Learn allows members to communicate using either an ad-hoc or infrastructural network by permitting the team members to be spread over a relatively larger geographical area when compared to some of the more traditional approaches towards collaborative learning. *Co-Learn* also allows for live streaming of audio/video files directly from the members to the presentation channel, which could either be the team leader's PDA or a desktop computer. During the streaming process, should the network drop, *Co-Learn* automatically resorts to recording mode. At this point in time, *Co-Learn* will allow the team leader to choose between waiting for the network to resume or switch to the next queued request. Once the network is resumed and if

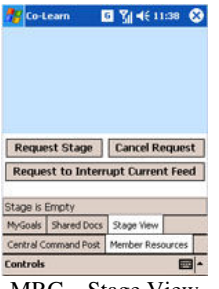
the member still 'has stage', *Co-Learn* will send the recorded message to the presentation channel and let the member resume the live feed. If the member has lost the stage focus, a queue request can be sent to the team leader for approval.

At the time this paper was being written, *Co-Learn* was still in its initial phases of design and development. We have developed a non-functional rapid prototype to organize our ideas and provide us a skeletal structure for *Co-Learn*. The following images have been taken from *Co-Learn's* rapid prototype.



MRC – Goals Display

MRC – Shared Documents



MRC – Stage View

Fig. 2. Member Resource Center – Screen Shots

5 *Co-Learn* Research Challenges

One of the biggest challenges for *Co-Learn* is that it must live within the constricted hardware environment, such as low processing power, minimal storage, battery operation, and very little graphics processing. These features should work at peak efficiency without taking a toll on battery power. The proposed soft real-time could become relatively large due to possible network delays. Since the devices are meant to be highly mobile, the network signal could rapidly deplete, which could cause delays or severely affect the quality of audio and video feeds. Also, we have yet to determine a way to encode and decode audio and video so that the processing speed and battery power are not severely affected.

In addition to design challenges due to constricted hardware, unique expectations of user-centered design introduce even more HCI (human computer interaction) research issues. A good user interface will facilitate entering the input and meaningfully displaying the output. Research should address personalization/ customization and data privacy needs. While most handheld devices have a small screen and no keyboard, a good user interface for *Co-Learn* must have a sense of plot and flow and be intuitive to use. It is also important for such a system to be highly secure, so no unauthorized collaboration takes place. Additionally, digital signatures should be embedded in each relayed piece of data.

Since *Co-Learn* is still in its initial phases of conception, there are quite a few foreseeable shortcomings. Although modern PDAs are not far behind, they have not quite caught up to their bigger, heavier counterparts. Hence the file transfer mechanism of large files might seem slow to those acquainted with the high-speed

networks on their desktop PCs. Also, during the file transfer operation there is no mechanism to resume the transfer should the network drop.

Co-Learn allows users to change permission settings of their own files; however, the settings affect everyone, i.e. users can either allow everyone or no one to view/edit a file. Although the initial concept of permission control will be limited, we plan to expand it to address more complex scenarios in the future.

6 An Illustrated Example

Let us take a look at a classroom situation where collaborative learning is at work. A broadcasting professor has divided the broadcasting class into three groups: Team A, Team B, and Team C. All students have handheld devices such as windows mobile or Palm OS-based PDA. The classroom also has two high-end graphics and video processing systems and video recording cameras. Each group is assigned the task of exploring the reaction of the student body to the recent decision to change the school's basketball team name and mascot within the class hour. Team A decides they will interview within the team and create a presentation. Teams B and C decide to dispatch six members to interview random students and capture their reactions. Team B uses standard video cameras, while Team C is equipped with *Co-Learn*.

While team A has a stronger presence in the classroom, they are limited in resources and are unable to bring 'out-of-classroom input'. Team B interviews a few people using traditional cameras; however, the team leader for Team B knows nothing about the location or progress of the team. The decision to provide additional out-of-class interviews suddenly becomes chaotic.

Team C, aided with *Co-Learn,* interviews people using their handheld device camera. The feed is relayed directly to either the processing station or to the team leader as an input. The team leader defines the goals and assigns them to the members (C3). The team leader can guide the team members by sending a list of interview questions (C1), (C4), (C5) and (C6). They can also monitor the length of an interview to avoid wasting time (C4) and (C6). While the 'in-class members' of Team C prepare the presentation, they create a list of suggestions which will help them to smooth the flow of the presentation, such as backdrop, ambiance, etc. and they then share these ideas with the team (C2). The in-class members are informed that a new document has been shared (C2) and they receive the requirements by visiting the MRC wall (C7). Field members start working towards implementing the suggestions, as long as they do not conflict with original goals. Input from different 'field members' is received in real time (C1). The team leader connects his/her handheld device to the projector, which allows them to not only do a live interview, but, using the voice module (C5) of *Co-Learn,* they are also able to take questions from the professor and other teams from directly outside the classroom.

In the above scenario, it can be seen how a system such as *Co-Learn* can aid students to effortlessly perform their tasks. *Co-Learn* allows students to focus on 'what to do' rather than 'how to do it'. In our *Co-Learn* system, we have attempted to achieve harmonious teamwork.

7 Evaluation

Co-Learn has been implemented using Microsoft .NET compact framework and is deployed on a combination of hardware and software, but it could be implemented on a J2ME platform. The targeted devices are Dell Axim 50v running Windows Mobile 2005, HP H6315 running Windows Mobile 2003 and Toshiba e740 running Windows Mobile 2002. Since one of the primary focal points of *Co-Learn* is user-centric design, it was deemed necessary that a rapid prototype be developed and tested. A survey was performed to gauge the user centricity of *Co-Learn*. The survey questions were presented to a group of twenty participants from varying backgrounds who ranged in age from 20 to 37 years and had varying technical knowledge. The questions were rated on a scale of 1 through 5, with 1 being least organized/important and 5 being well organized/most important.

The participants were given a short explanation of collaborative learning and then provided the rapid prototype to experiment with for a few minutes. After trying the device, they were asked to answer a set of questions based on their experience with *Co-Learn*. The results obtained were based on participant's experience with *Co-Learn* and their ability to understand and adapt to *Co-Learn* without any external help or prior exposure. Figure 4 displays the average scores obtained by the initial survey.

Our idea was to establish the user-centric design and collect input from people who have been in collaborative learning environments but had no prior exposure to *Co-Learn*. *Co-Learn* was determined to be fairly balanced between user-centric and system-centric design; we believe it has provided us with a good starting point. Based on the responses and suggestions obtained in this exercise, we will revisit *Co-Learn* to improve interface design and add more features. Also, this approach will help us present our next prototype to two different groups, those who have experienced *Co-Learn* in past and others who have not. In addition to survey questions, the participants were asked to rate each of the services/features provided by *Co-Learn*. The following chart represents the average scores on a scale of 1 to 5, with 1 being least important and 5 being most important.

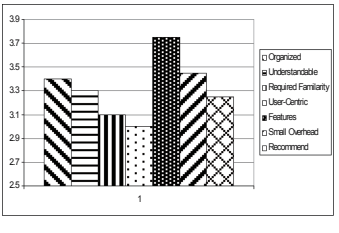


Fig. 3. Survey Results

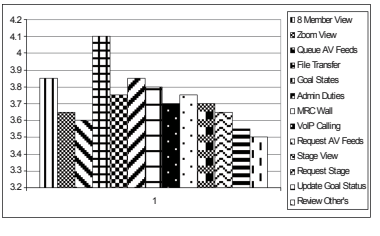


Fig. 4. Average Scores of Co-Learn features

8 Conclusions

Today education is an essential part of our life. With the incorporation of various technological advances into the modern classroom, we expect a richer learning experience. Although these tools [7, 9, and 10] are available, they are not always

focused on the user's needs. In this paper, we have proposed a system called *Co-Learn*. *Co-Learn* aims to aid students in their collaborative learning experiences, making it richer, and easier than ever to collaborate with each other in team settings. We have developed the first prototype of the *Co-Learn*, in order to gauge our understanding of user-centric design techniques. Using the results obtained from our initial survey we will re-visit *Co-Learn* and make improvements and enhancements. The design phase of *Co-Learn* will focus on harnessing the power of handheld devices and presenting it to the end users in a simple but effective way. We will re-evaluate *Co-Learn* by presenting the new prototype to the same set of users as in our initial evaluation, and to those who have no experience with *Co-Learn*. This will allow us to re-evaluate *Co-Learn's* overall ease of use as seen by first time users and by those who have been exposed to *Co-Learn* in the past. We also plan to extend the audio and video channels to more than eight users and address the access control and security features of our current design. In the future, we plan to improve and test *Co-Learn* in a real classroom.

References

1. Davis, M. et al.: MMM2: Mobile Media Metadata for Media Sharing. MM'05. Singapore, pp. 267–268 (2005)
2. US Department of Education – National Center for Educational Statistics - http://nces.ed.gov/programs/coe/2006/section1/table.asp?tableID=443
3. Gartner: PDA shipments reached record high in Q2 of 2006 (2006) http://www.emsnow.com/npps/story.cfm?ID=21090
4. Flaschbart, J., Franklin, D., Hammond, K.: Improving Human Computer Interaction in a classroom environment. In: Proceedings of the 5th international conference on intelligent user interfaces, pp. 86–93 (2000)
5. Gokhale, A.: Collaborative Learning Enhances Critical Thinking. Journal of technology Education. vol. 7(1) (1995)
6. Holzinger, A.: Fundamentals of Human Computer Interaction (HCI) for e-learning. Innovative concepts for teaching informatics, pp. 102–112 (2005)
7. Holzinger, A., Nischelwitzer, K., Kickmeier, M.: Pervasive E-Education Supports Life Long Learning: Some Examples of X-Media Learning Objects. In: Proceedings of 10th IACEE World conference on Continuing Engineering Education (2006)
8. Holzinger, A., Nischelwitzer, K., Meisenberger, M.: Lifelong-learning support by M-Learning: example scenarios http://www.elearnmag.org/subpage.cfm?section=research&article=6-1
9. Yau, S., Gupta, S., Karim, F., Ahamed, S., Wang, Y., Wang, B.: Smart Classroom: Enhancing Collaborative Learning Using Pervasive Computing Technology. In: Proceedings of American Society of Engineering Education 2003 Annual Conference (2003)
10. Kildare, R., Williams, R., Hartnett, J.: An Online tool for learning collaboration and learning while collaborating. In: Proceedings of the 8th Australian conference on Computing education, vol. 52, pp. 101–108 (2006)
11. Rummel, N., Spada, H., Hauser, S.: Learning to collaborate in a computer-mediated setting: observing a model beats learning from being scripted. In: Proceedings of the 7th international conference on learning sciences, pp. 634–640 (2006)

Effects of Group Composition in Collaborative Learning of EFL Writing

Siew-Rong Wu

Center for General Education, National Yang-Ming University, 155 Li-Nong Road, Section 2, Peito District, Taipei 112, TAIWAN
srwu@ym.edu.tw

Abstract. This study was conducted to compare two types of group composition and their effects on collaborative learning of EFL writing in PBL lessons. Participants in the experiments were 60 students taking the researcher's basic writing classes in the fall semester, from September 2003 to January 2004. Outcomes of bimonthly automatic online writing assessments using *My Access* showed that group composition did affect writing learning outcomes. Small groups of learners with even distribution of levels of writing proficiency did outperform the groups with randomly mixed learners of varied levels of English writing proficiency. This finding indicated that when less capable learners collaborated with more capable learners in small group learning activities, the learning outcomes could be greatly enhanced. This implies that when taking the PBL approach in learning to write, groups composed of learners with evenly distributed levels of writing proficiency may achieve much better learning outcomes than randomly formed groups.

Keywords: group composition, automatic online writing assessment, Zone of Proximal Development, collaborative learning, PBL, ill-structured problems, knowledge facilitator.

1 Introduction

This study centers on the rationale of Vygotsky's notion of the Zone of Proximal Development (ZPD), in which less capable learners can learn from more capable learners in the social context by interacting with each other. Collaboration among learners will promote growth, and learners of similar ages are likely to be operating within each other's proximal zones of development [3]. Collaborative learning is beneficial in facilitating learning because through active participation in meaning-making, learners would be able to engage more deeply in their own knowledge construction as they integrate the new knowledge into their own schema and present it meaningfully.

However, despite all these endeavors in the exploration of knowledge management, few studies have yet examined the difficulties EFL learners encounter in collaborative online learning of writing. Outcomes of a preliminary study [4] show that in a conventional Taiwanese university classroom, peer conferencing is beneficial

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 843–849, 2007.
© Springer-Verlag Berlin Heidelberg 2007

to the EFL university freshmen's learning of scientific writing in English, especially in the enhancement of critical thinking and scientific writing abilities.

Based on the above outcomes, a question arises. In collaborative online learning of EFL writing, what kind of group composition will bring about better learning effects – randomly formed groups or groups of evenly distributed levels of writing proficiency? To answer this question, theme-based instruction and learner-centered collaborative online learning activities designed within the framework of PBL were implemented in the experiments. The findings would serve as an essential reference for the teaching of EFL writing in which the PBL approach is taken.

2 Participants, Group Compositions, PBL Lessons, Automatic Writing Assessment, and Discussions

2.1 Participants

The participants in this study were 60 undergraduates (44 males and 16 females; most of them were aged 18 to 19) taking the researcher's three elective Internet-assisted writing classes at a medical university in northern Taiwan in the fall semester of 2003, starting from September 2003 to January 2004. The maximum number of learners in each of these classes was 20. The undergraduates at this university are required to take two to four credit hours of foreign languages, i.e., English, German, or Japanese, in their freshman year, and they had all received six years of previous English training in high schools and middle schools. However, their levels of writing proficiency still varied greatly.

2.2 Group Compositions and the Pretest

The pretest serving as the baseline for comparing learning outcomes of PBL lessons 1 and 2 was given right at the beginning of the semester. In the pretest, all students took the first automatic writing assessment using *My Access*. In PBL lesson 1, participants formed small groups of 5 randomly. Posttest 1 was performed immediately after PBL lesson 1 was completed. However, group composition was changed in PBL lesson 2. This time, each group was assigned participants with evenly distributed automatic writing assessment scores. Posttest 2 was performed after PBL lesson 2 was completed.

2.3 PBL Lessons

Two PBL lessons were designed to engage and motivate the participants. PBL lesson one was an ill-structured real medical problem that the researcher just encountered. Through recursive scientific inquiry processes, these learners collaborated in problem-solving activities such as hypothesizing, information search, discussions, and revisions of hypotheses. In PBL lesson two, learners were asked to explore the meaning of *The Lord of the Rings, Part III: The return of the King* and the descriptive skills of the writer.

2.4 Automatic Writing Assessments

An automatic writing assessment tool, *My Access*, Version 4, developed by the Vantage Learning, Inc., USA, was used for bimonthly writing assessments. Each of the learners' essays was automatically analyzed in five domains, i.e., focus, development, organization, style, and conventions. A 6-point scale based on scoring rubrics was used in scoring, with 1 being the lowest score and 6 the highest. A holistic score with all five domain scores was produced several seconds after online submission of essays. Meanwhile, diagnostic analyses were also provided automatically along with the holistic and domain scores. Two prompts were used for automatic writing assessment after each PBL lesson. Each essay was written online and submitted within 30 minutes.

2.5 Peer Conferences and Student-Teacher Conferences

To facilitate learning in the PBL lessons, the teacher acted as an encouraging knowledge facilitator and refrained from controlling the content of discussions and information search. Guidance was provided in the form of questions to elicit critical thinking and scaffold knowledge construction. These interactions were carried out in the discussion forums of the researcher's teaching web site. Learners also gave critical feedback to each other. Peer conferences, face-to-face student-teacher conferences, as well as online discussions were done all through the learning processes in both PBL lessons.

3 Results

3.1 Scores of Automatic Writing Assessments

Table 1 indicates a fully positive correlation between the holistic scores and the domain scores learners obtained in each automatic writing assessment and prompt. The results were obtained by running Pearson's correlation tests ($P < 0.001$).

Table 1. Holistic scores vs. five domain scores of assessments for the two prompts

D	S		p			p
	S	S S	S S	S	S S	S S
Focus	3.2	3.3	3.8	3.2	3.2	3.6
Development	2.8	3	3.4	3	3	3.3
Organization	2.7	2.9	3.3	2.8	3	3.2
Style	3.3	3.3	4	3.4	3.4	3.6
Conventions	3.3	3.3	3.9	3.4	3.3	3.6
Holistic score	3.2	3.3	3.9	3.2	3.3	3.7

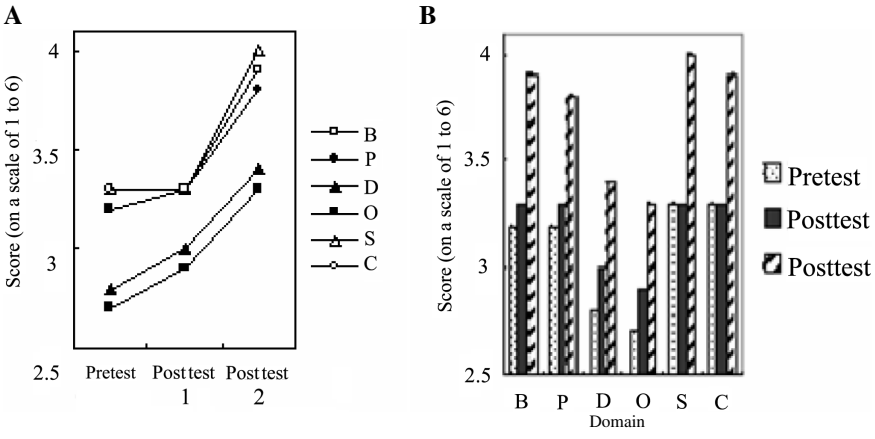


Fig. 1. Scores of all students' automatic writing assessments for prompt 1: "Effects of Technology" (n=60). H stands for Holistic Score; F for the domain of Focus; D for Development; O for Organization; S for Style; C for Conventions. (A) Students' average holistic scores vs. the average domain scores. (B) Students' average scores vs. those of the pretest, posttests 1 and 2.

Fig. 1A shows that there is a huge gap between the domain scores of style, conventions, focus and the domain scores of development and organization all through the three assessments. On the other hand, Fig. 1B shows that little progress was made between pretest and posttest 1, while huge progress was made between posttest 1 and 2.

Fig. 2A indicates that there is a huge gap between the domain scores of focus, conventions, and style, and the domain scores of development and organization in the final assessment, and that organization was the most difficult writing skill to master. On the other hand, Fig. 2B shows that relatively less progress was found between average scores of pretest and posttest 1 than the scores between posttests 1 and 2.

Two one-way ANOVA tests were performed to examine whether there was significant difference between the three holistic scores in the three assessments using the two prompts. Results shown in Fig. 3 indicate that the three holistic scores for prompts 1 and 2 were significantly different (F=28.574; df=2,58; P 0.001) (F=13.267; df=2,58; P 0.001).

Individual pairwise comparisons of these two prompts show that there was no significant difference between pretest and posttest 1 (I-J=0.083; 0.0) (I-J=0.1; 0.0) while significant difference was found between pretest and posttest 2 (I-J=0.7; 0.001) (I-J=0.483; 0.001) as well as between posttests 1 and 2 (I-J=0.617; 0.001) (I-J=0.383; 0.001) for both prompts.

A two-way ANOVA test was performed to test if the use of different prompts and group compositions had main effect on automatic writing assessment results and to examine the interactions between these two factors. Results displayed in Fig. 4

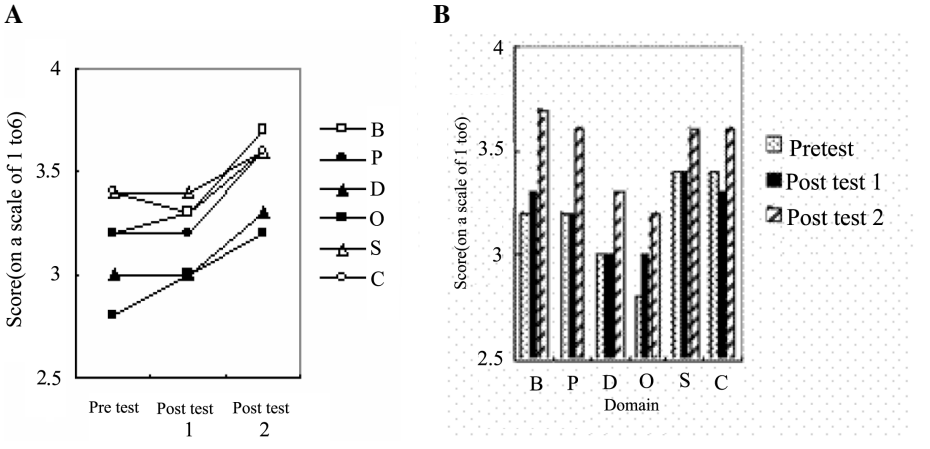


Fig. 2. Scores of all students' automatic writing assessments for prompt 2: "Mandatory Recycling" (n=60). (A) Students' average scores vs. the average domain scores. (B) Students' average scores vs. those of the pretest, posttests 1 and 2.

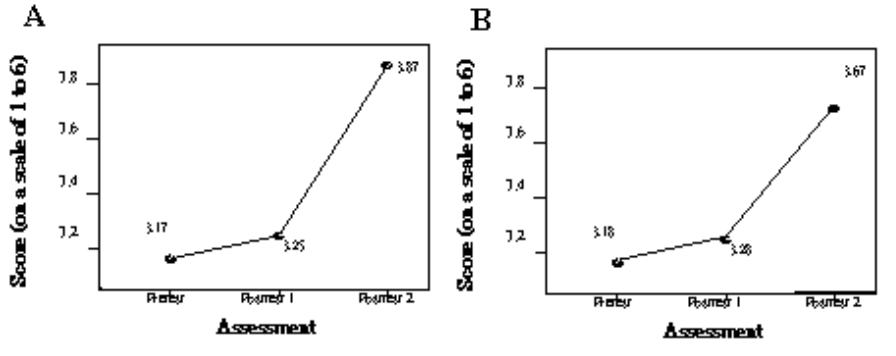


Fig. 3. Holistic scores for prompts 1 and 2. (A) All learners' holistic scores for prompt 1. (B) All learners' holistic scores for prompt 2.

showed that prompts did not produce main effect on the automatic writing assessment results significantly (F=1.578; df=1,59; $P > 0.0$. That is, within the same group of 60 learners, the use of any of the two prompts did not significantly affect the scores of automatic writing assessments. On the other hand, the difference of group composition did significantly produce main effect on the automatic writing assessment results (F=29.5; df=1,59; P 0.001). Besides, interactions were also found between the prompts and different group compositions (F=4.308; df=1,59; P 0.05).

Subsequently, two t-tests were performed to examine if there was a significant difference between the two prompts in posttests 1 and 2. Results showed that the two prompts in posttest 1 did not produce significant difference (t=-0.375; df=59;

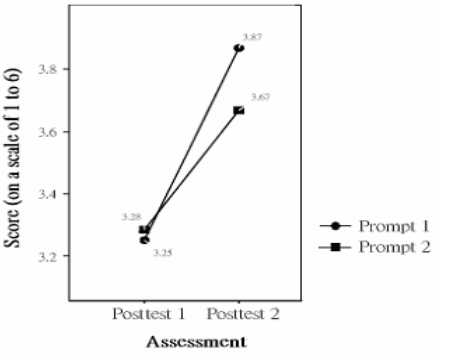


Fig. 4. Holistic scores of prompts 1 and 2 in posttests 1 and 2

P 0.05; two-tailed), but significant differences were found between the two posttest scores (t=2.352; ; 0.0 ; two-tailed . The scores of posttest 1 were significantly higher than those of the posttest 2.

3.2 Students' Reflections on Learning

Most of the EFL learners in Taiwan were used to the conventional passive ways of learning. Therefore, many of the participants in this study were not used to this learner-centered approach in which they were required to do lots of self-learning, even after class. The reluctant and less proficient learners complained about the heavy load of learning. Conversely, the more proficient learners reflected that they enjoyed this type of collaborative online learning within the PBL framework because it made them learn actively, and that through collaboration and conferencing with peers, they have learned so much more than in the conventional classrooms. They reflected that as they helped their peers, they got to learn more in the meanwhile.

4 Discussions

Score analyses of the results in Fig. 1 and 2 indicated that group compositions might affect learning outcomes, as we can see from the huge gap between the domain scores of focus, style, and conventions vs. development and organization for prompt 1 all through the pretest, posttests 1 and 2. Besides, little progress was found between the scores of pretest and posttest 1, but a great leap was found in the scores of the final assessment, as compared to those of the earlier two assessments. For prompt 2, we can also find a huge gap between the domain scores of focus, conventions, and style vs. the domain scores of development and organization in the final assessment. Relatively less progress was found between average scores of pretest and posttest 1 than the scores between posttest 1 and 2.

The experiment results in this study revealed that learners with even distribution of levels of writing proficiency in a group did outperform those in groups in which learners' levels of writing proficiency were randomly mixed. This finding indicated

that when less capable learners collaborated with more capable learners in learning activities in small groups, the learning outcomes could be enhanced. This was probably done through a scaffolding process [1], in which more capable learners engage in learning activities with less capable learners. To conclude, when learners were grouped randomly, using different prompts did not lead to significant difference in learners' performance in automatic writing assessments. In contrast, when learners were grouped with even distribution of levels of writing proficiency, they did perform significantly better in prompt 1 than in prompt 2.

Furthermore, when adults teach children, the appropriation process is mutual, and the interactions are beneficial to both sides. Meaning construction is thus collaboratively achieved [2]. The more capable learners' reflections also indicated that as they helped their peers and gave critical feedback to them, their writing proficiencies were also enhanced. This implies that when taking the PBL approach in learning to write, groups composed of learners of evenly distributed levels of writing proficiency, in which there were both more capable and less capable learners, would achieve much better learning outcomes than randomly formed groups. This is important pedagogically.

However, there might be interactions between the learning effects of lessons 1 and 2, which might also influence the assessment outcomes. Meanwhile, as learners became more adapted to the PBL approach and could enjoy collaborative learning more, they might hence perform better in writing. Learners' getting used to the automatic writing assessment tool might be another important factor that influenced the assessment results. Therefore, further studies are needed to explore this issue more in depth before any absolute conclusion can be made.

Acknowledgements. This study was funded by a grant (no. NSC 92-2520-S-010-001) from the National Science Council of Taiwan, Republic of China and a grant for the academic year of 2003 from the Distance Education Committee of National Yang-Ming University.

References

1. Bruner, J.: From communication to language: A psychological perspective. Cognition 3, 255–287 (1975)
2. Newman, D., Griffin, P., cole, M.: The construction zone: Working for cognitive change in school. Cambridge University Press, New York (1989)
3. Vygotsky, L.: Mind in society: The development of higher psychological processes. Harvard University Press, Cambridge, MA (1978)
4. Wu, S.R.: Peer conferencing in a university scientific writing class, University Language Centres: Forging the Learning Environments of the Future. In: Satchell, R., Chenik, N. (eds.) Chapter 3: Language for Specific Purposes/Language for Academic Purposes, Confederation Europeene des Centres de Langues de l'Enseignement Superieur. Paris, France pp. 223–236 (2004) ISBN 2-9522811-0-6

E- Health System for Coagulation Function Management by Elderly People

Noemi Biterman[1], Eyal Lerner[1], and Haim Bitterman[2]

[1] Industrial Design, Faculty of Architecture and Town Planning
[2] Faculty of Medicine, Technion, Israel institute of Technology, Haifa 32000, Israel

Abstract. E- Health is a developing area of major social, medical and economic importance especially for the elderly population and citizens of remote areas. *Our objectives* were to identify visualization methods for a patient-oriented system of collection, storage, and retrieval of coagulation function data, The research group included 25 elderly (72.2± 5.5 years) and 25 young participants (30.4±4.9 years).The participants completed tasks based on different visualization models for data entry and follow-up of clinical information, in three experimental websites equipped with hidden tracking programs. We followed functional parameters (time, acuity), subjective parameters (preference, satisfaction) and physiological parameters (pulse, skin temperature, sweating, respiratory rate, and muscle tension).Time for task completion was significantly longer in elderly compared to younger participants in all experimental websites, without significant differences in accuracy. Yet, in specific tasks the elderly performed better than young participants. Specific suggestions for data entry and data visualization are presented.

Keywords: E-care, interactive healthcare services, data visualization, patient-oriented systems, telemedicine, design methods.

1 Introduction

E- Health is a developing area of major social, medical and economic importance [1]. Interactive Health Services are expected to improve the chain of treatment, decrease healthcare expenses, enhance quality and efficiency of healthcare, and increase the fairness and equality of distribution of medical services, mostly for the elderly population and citizens of remote areas [2, 3]. Interactive Health Services will be an integrated part of the "knowledge society" in the next century that will also be the era of the aging population [4].

There are evidences that interactive health services, by which the patient continuously takes care of his own data, can improve clinical outcome of chronic diseases (such as heart failure, hypertension, diabetics), reduce the number of clinic visits, allow earlier discharge from hospital and increase patient's satisfaction [2, 5]. However, benefits emerging from the new technology of the "Knowledge Society" depend upon its adaptation to the needs and possibilities, and to the willingness of the elderly to use it [6].

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 853–859, 2007.
© Springer-Verlag Berlin Heidelberg 2007

There are only few controlled studies, which are directed towards exploring and discussing the specific adaptability of *medical* human-interactive computer interfaces (web-based) for old people.

Our objectives were to identify visualization methods for patient-oriented systems of self-monitoring, collection, storage, and retrieval of medical data for elderly people. As a case study we selected a system for self control of coagulation data in patients with common medical conditions (e.g. heart rate disturbances, prolonged immobilization or artificial heart valves) that require chronic anticoagulant treatment with warfarin derivatives (e.g. Coumadin) to prevent and treat thrombosis. The dose of the anticoagulant drug is closely scrutinized by performance of frequent blood tests (INR test) and maintaining it within a narrow therapeutic range by changes in the dose of the drug [7]. A combination of the currently available home device for self performance of INR blood tests, education of patients on drug dosage adjustment, and an efficient computer based self-monitoring system will free the patients from the traditional tedious and often impossible routine of frequent clinic visits for performance of the test and for medical instructions.

2 Methods

2.1 Participants

The study included 25 elderly people (72.2 ± 5.5 years) and a control group of 25 young people (30.4+4.9 years) having previous experience with internet and computer applications. All participants were equipped with miniature sensors for continuous recording of physiological parameters (fig 1).

Fig. 1. A participant equipped with miniature sensors for continuous recording of physiological parameters indicating stress and anxiety: heart rate, sweat and skin temperature, muscle tension and respiratory rate) (AR7- Stress Management Monitor"; Atlas Res. Ltd., Israel)

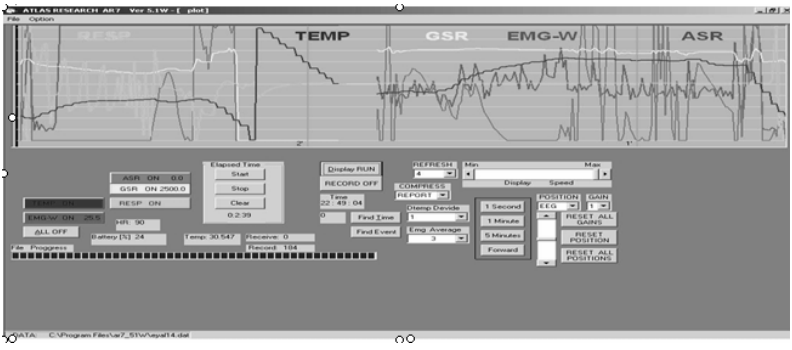


Fig. 2. A typical output of the physiological monitoring system (ATLAS)

The measured parameters, representing stress and anxiety were : heart rate, skin sweat, muscle tension (Electromyography), skin temperature and respiration (fig 2), using "AR7- Stress Management Monitor" (Atlas Res.Ltd., Israel).

2.2 Experimental Websites

Three experimental websites based on different visualization models for data entry and follow-up of personal clinical information were constructed and equipped with hidden tracking programs, to fully record the participants' steps and navigation routes.

In the first part of the experiment the participants were asked to introduce the results of the coagulation test (INR) they obtained during that day into three sites via different input methods: free text fields, menus, and visual buttons. In the second part the participants viewed reports of their data displayed by three methods: table, graph, or calendar format, and were asked to answer six questions based on common clinical scenarios.

The tasks included identification of a specific parameter, detection of irregular values and trend recognition. The three web sites were similar in their level of complexity and color-scheme, and differentiated only by their user's models and/or design display.

2.3 Experimental Parameters

Three groups of parameters were measured: Functional parameters (total and relative time for task completion, errors, number of extra keystroke and navigation between reports); Physiological parameters (heart rate, sweating, skin temperature, respiratory rate and electromyography) and Subjective parameters (preferences and satisfaction).

2.4 Statistical Analysis

For descriptive statistics, continuous variables were expressed as mean ± SD, and categorized variables as percentages. The Chi-square (X^2) or Fisher's exact test were used for comparison of categorized values and t-test for normally distributed continuous data.

The different interfaces were compared within each age group and between the two age groups for each web configuration. Comparison was carried out using repetitive analysis of variance tests, and ad hoc Duncan test.

Comparison between age groups was done by unpaired t-test.

A value of $P < 0.05$ was considered significant.

3 Results

3.1 Differences Between Age Groups

Total time for task completion (data input and data follow-up) was significantly longer in elderly compared to younger participants in all experimental websites ($P<0.0001$; $F=119.8$ and $P<0.0001$; $F=64.07$, respectively), as can be seen in figures 3 and 4, respectively. No significant differences were found in accuracy between elderly and young participants as reflected in the number of errors.

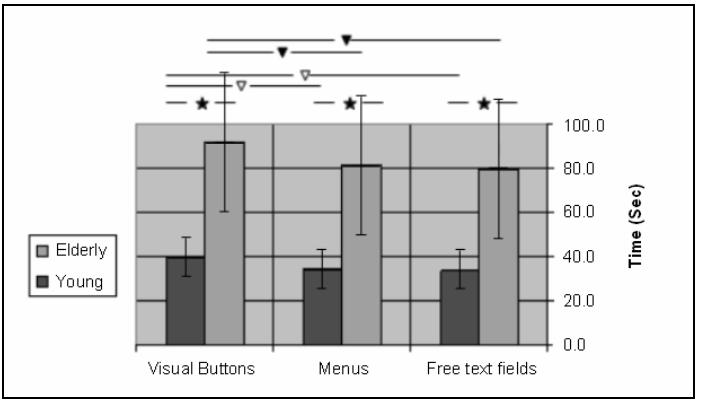


Fig. 3. Total time for input of clinical data, in the three websites, by both age groups (n=25 for each age group). * $P<0.05$ in t-test, comparison between age groups. ▽▼ $P<0.05$ in Duncan test, comparison between websites in each age group.

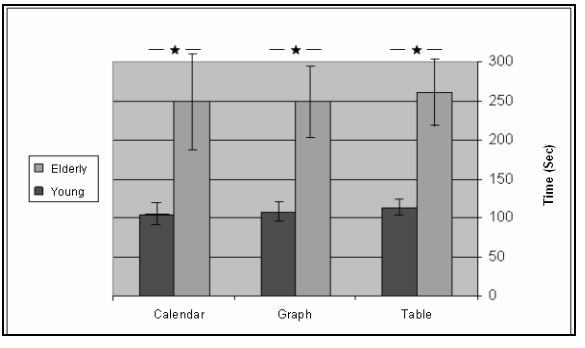


Fig. 4. Total time input of follow up of clinical data, in the three websites, by both age groups (n=25 for each age group). * $P<0.05$ in t-test, comparison between age groups.

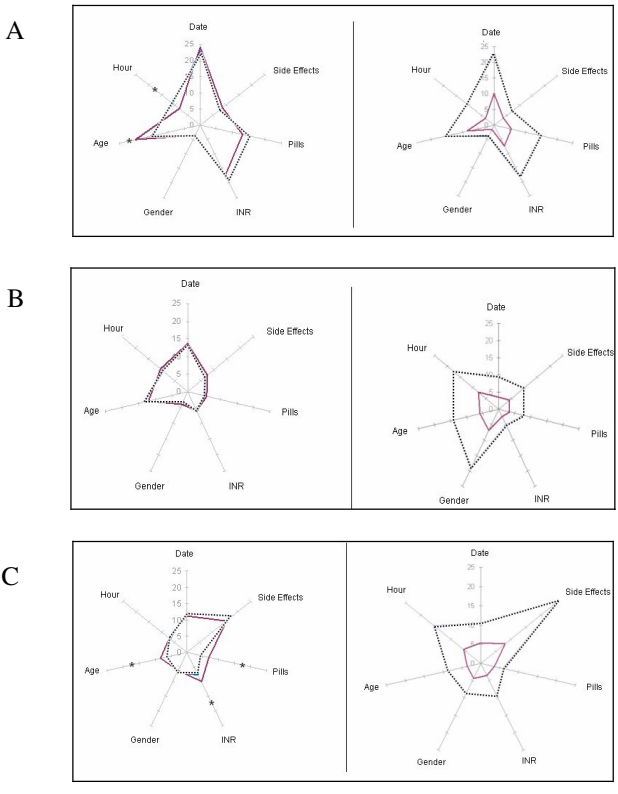


Fig. 5. Absolute (right side of picture) and relative time (left side) for data insertion in both age groups. A- visual buttons, B- Menus, C- Free text field . Continuous line (-) young participants, dashed line (---) elderly participants. * P<0.05 in t-test, comparison between age groups.

In most parameters there were no statistical differences between age groups in the relative time spend in each task. Yet, in specific tasks the elderly performed better than young participants (e.g insertion of "age" in the visual buttons website P=0.01 ; F=7.17), while younger participants performed faster in inserting "hour" in the visual buttons website (P=0.0008 ; F=12.69) (fig 5). There were no significant differences in changes in physiological parameters between the two age groups.

3.2 Difference Between Websites

Comparison between websites revealed that the longest total time for inserting data was measured in the visual buttons website compared to free text and menus in both age groups (P<0.0009 ; F=7.6) (fig 3). Table display was the best design for locating a specific clinical value, and the graph presentation was the worst (Duncan, P<0.05) . For locating abnormal values, the graph was the best design and the table was the worst visualization method (Duncan, P<0.05) in both age groups. No correlation was found between the number of errors and web site configuration in both age groups.

Significant increases in heart rate and skin sweat were measured in the free text configuration (for data insertion) compared to the other two websites in both age groups.

4 Discussion

Our results demonstrate that older adults can perform Internet tasks as well as, though slower than, young people, as has already been shown previously [8, 9]. The absence of differences in physiological parameters between the two age groups supports our claim the diversity in internet performance is not stress related but represents age-related mechanisms.

The use of relative time for task completion indicates that there are specific tasks that are performed better by elderly, while others are done better by young participants. Therefore we suggest that each website should be checked experimentally by both age groups to support both groups' needs.

A tabular display was preferred by both age groups for identification of specific clinical information. A graphic display was preferred when looking for an exceptional clinical value. No differences were found for trend recognition between websites. Specific configurations should therefore be used for different tasks.

No correlation (and even an occasional negative correlation) was found between satisfaction and preferences rankings and performance scores, suggesting that the users do not always know what is best for them [10]. This discrepancy between function and preferences emphasizes the need for using objective measures.

The use of continuous monitoring of physiological parameters adds objective measures of stress for evaluating web functions. Yet, a better understanding of the correlation between physiological parameters, functional HCI measures, and subjective satisfaction criteria are still needed before we can introduce it as a conventional assessment tool [11] .

Interactive computer interfaces should be fully accessible not only for the elderly, but for people with functional limitations related to disabilities. One should also remember that circumstances and external situations (such as background noise, dim light, stress, unsteady and moving environment, etc.) place ordinary people in a position comparable to that of a disabled person [12,13].

We expect to establish some general models and concepts for the design of interactive medical computer system interfaces compatible for elderly people, that could be modified and tailored (personalized) according to different diseases, changing needs of the patients, and possibly gender.

References

1. Gomez, E.J., Del Pozo, F., Arredondo, E.: Telemedicine: A new model of health care. J. Healthcare Technology and management 3(4), 374–389 (1999)
2. Rione, R., Ohinmaa, A., Hailey, D.: Assessing Telemedicine: A Systematic Review of the Literature. CMAJ 165, 765–771 (1999)
3. Morrell, R.W.: Older adults, Health information and the World Wide Web. Lawrence Erlbaum Associates, Mahwah, NJ (2002)

4. Drucker, P.: The next society. A survey of the near future. The Economist, 361, 3S–22S (2002)
5. Chae, Y.M., Heon Lee, J., Hee Ho, S., Ja Kim, H., Hong Jun, K., Uk Wan, J.: Patient satisfaction with telemedicine in home health services for the elderly. Int. J. Med. Inform 61, 167–173 (2001)
6. Bitterman, N., Shalev, I.: The Silver Surfer; Internet for Senior Citizens. Ergonomics in Design 12, 24–28 (2004)
7. Deitcher, S,R.: Antiplatelet, anticoagulant, and fibrinolytic therapy. In: Kasper, D.L., et al. (ed.) Harrison's Principles of Internal Medicine, 16th edn. pp. 687–693. McGraw Hill, New York (2005)
8. Mead, S.E., Spaulding, V.A., Sit, R.A., Meyer, B., Walker, N.: Effects of age and training on world wide web navigation strategies. In: Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting, pp. 152–156 (1997)
9. Morrell, R.W., Dailey, S.R., Feldman, C., Mayhorn, C.B., Echt, K.V.: Older adults and information technology: A Compendium Of Scientific Research And Web Site Accessibility Guidelines. Washington, D.C.: National Institute on Aging (2001)
10. Ward, R.D., Marsden, P.H: Physiological responses to different web pages design. International Journal of Human-Computer Studies 59, 199–212 (2001)
11. Andre, D.A., Wickens, C.D.: When users want what's NOT best for them. Ergonomics in Design 7, 10–14 (1995)
12. Vanderheiden, G.: Design for people with functional limitations due to disability, aging, or circumstances. In: Salvendy, G. (ed.) Handbook of human factors and ergonomics, pp. 2010–2052. John Wiley & Sons, New York (1997)
13. Murphy, D.R., Craik, F.I.M., Li, K.Z.H., Schneider, B.A.: Comparing the effects of aging and background noise on short-term memory performance. Psychology and Aging 15, 323–334 (2000)

Feed the Dragon Wisely: Designing for Childhood Awareness as a Means of Lifelong Obesity Prevention

Shunying Blevis[1], Jeffrey Bardzell[1], and Nancy Wroblewski[2]

[1] School of Informatics, Indiana University Bloomington, IN 47408
[2] Monroe county YMCA, 2125 S Highland Ave, Bloomington, IN 47401
san@indian.edu, jbardzel@indiana.edu,
vacationfamily@insightbb.com

Abstract. Obesity is becoming a global epidemic. Obesity education, especially childhood obesity education, has become an issue for the public health community agenda in a lot of countries. In this paper we will report on an interactive educational toy we designed and prototyped for obesity education. We propose design refinements and marketing strategies to make the toy available to educate children more widely and effectively.

Keywords: Obesity education, Childhood obesity, Interactive educational toy, Tangible interface.

1 Introduction

With economic development, modernization, and urbanization, the problems of being overweight and of obesity are becoming a global epidemic [1]. Obesity education, especially childhood obesity education, has become intrinsic to the public health community agenda in a lot of countries.

Although compelling media such as television and video games can contribute to the obesity problem [7, 8], these media can also be effective educational tools for obesity prevention education, especially among young children. In this paper we will report on an interactive educational toy we designed and prototyped for obesity education. We propose the design refinements and marketing strategies to make the toy available to educate children more widely and effectively.

2 Background

Obesity is becoming a global epidemic, with more than 1.6 billion adults overweight and at least 400 million of them clinically obese globally in 2005. The obesity rate continues to rise quickly in both developed and developing countries [1], and the consequences are high. Being overweight or obese increases the risk of many diseases and health conditions, such as hypertension, stroke, diabetes, etc. [2], making it one of the major health problems of 21st century.

The causes of obesity are usually related to lifestyle, specifically unhealthy diet and physical inactivity [3, 4], making obesity, and the conditions it causes, largely

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 860–868, 2007.
© Springer-Verlag Berlin Heidelberg 2007

preventable. Promoting healthy eating and physical activity is a sustainable and effective strategy to prevent obesity and promote health [4, 5].

Recent data suggests strategies we can adopt to prevent obesity, and one of them is to educate children about eating, health, and obesity. The latest projections of the World Health Organization [WHO] indicate that at least 20 million children under the age of 5 years were overweight globally in 2005 [1]. According to the Center for Disease Control (CDC) in the United States, the percentage of children who are overweight has more than doubled since 1980 [2]. Overweight young people are more likely than children of normal weight to become overweight or obese adults [6]. So the choices we make early in life will affect our health later in life and it is very important to teach children to choose healthier life styles when they are young.

As a result, obesity education, especially childhood obesity education, has become a major part of the public health community agenda in a lot of countries. For example, in May 2004, WHO has adopted the "Global Strategy on Diet, Physical Activity and Health" [4]. In the United States, the CDC has supported numerous national, state and local programs to address childhood obesity problem. [6]

3 Research

In designing an educational toy for teaching children about eating well and making choices that will not lead to obesity, we focused our research in three areas: existing programs that target obesity education, so we knew what was already out there; childhood learning, so that our design was developmentally appropriate; and the intersection of children and computers, so we could understand the kinds of interactions most likely to appeal to children.

3.1 Existing Programs/Products Targeting at Childhood Obesity Education

Numerous programs and media already exist that target obesity education for children, and we set out to identify the structures and activities of these programs, their targeted learning objectives, and to observe children in them.

YMCA Health class. We went to observe a health class for pre-school children in YMCA two times. The class was divided into two parts; in the first part, an instructor offered a lecture focusing on the new food pyramid; and in the second part, children participated in activities, such as feeding a puppet bird or preparing a nutritious lunch. We observed that children are generally more engaged in the activities than the lectures, and so we designed our approach around activity, direct manipulation of tangible objects, rather than abstract, symbolic discourses (such as the food pyramid and verbal explanations of it).

TV shows. One of our inspirations was Cookie Monster, a popular character on the children's television show *Sesame Street*. In previous generations, this character used to sing "C is for cookie, good enough for me." More recently, Cookie Monster advocates healthy eating by saying "A Cookie is a Sometimes Food" and beginning with a "health tip" about healthy foods and physical activity in each episode of the show's new series [12].

That *Sesame Street* chose to rework Cookie Monster's message, rather than phase him out in favor of a character more obviously appropriate, suggests that his enduring appeal, which has transcended decades of children, is worth more than his putative message. This suggests the efficacy of persistent characters in children's education and entertainment, and accordingly our design focused on developing a persistent character whom children would recognize and identify with.

Computer Games. Several edutainment computer games have been developed to address the issue of obesity. Of course, the obvious problem with these is that they encourage lethargy by requiring children to sit still in chairs to use them. Nonetheless, their interaction design was still of interest to us, prompting us to ask whether we could design similar interactions while encouraging children to act.

a) Feed my FURBY online is an online computer game, in which children can choose and feed food to FURBY and FURBY's shape will change according to its diet. [11]
b) MyPyramid Blast Off Game is an interactive computer game where children can reach Planet Power by fueling their rocket with food and physical activity. "Fuel" tanks for each food group help children keep track of how their choices fit into Food Pyramid. [10]
c) Nutrition and Kids Adventures is a 3D animation game, which takes children on a fun journey through fresh gardens and pyramids of food with characters like Banana Bana, Sidekick Leaf, Bear and the professor. [9]

One of the main benefits of these applications is that they help children visualize the consequences of their choices, which is much harder without using interactive media.

Books. We also looked at books that teach health and nutrition. One example is *Oliver's milk shake*, a children's series in which Oliver, the main character, teaches kids about nutrition and recipes. Children's books are especially effective at situating appealing characters in simple narratives, which, in addition to their intrinsic interest, also help clarify causal relationships between choices and consequences.

In all these examples, we commonly see characters in situations that reveal relationships among options, decisions, and consequences. Each medium has its own strengths in terms of clarifying these relationships; for example, computer games are good at interactively visualizing consequences, while children's books excel with clarifying causal chains.

3.2 Childhood Learning

As adults, we can critique different media as platforms for the kinds of messages we want to communicate and experiences we want to create. But, of course, if the design that arises of this critique is not developmentally appropriate for the children whom we are trying teach, then it is all for naught.

Social Learning. Social learning plays a very important role in the early childhood. Children acquire and develop behaviors by observing and then imitating other social models. As Bandura writes, "most human behavior is learned observationally through modeling: from observing others one forms an idea of how new behaviors are performed, and on later occasions this coded information serves as a guide for action." [13] Children can adopt modeled behaviors (including eating behavior) from adults, their peers and even the cartoon heroes they admire.

Activity theory and constructive learning. Conscious learning emerges from activities and constructive learning environments are created in classes, computer-based media and so on. Constructivist educational theorists Jonassen and Rohrer-Murphy write that "constructive learning environments should attempt to replicate the activity structures, tools and sign systems, socio-cultural rules, and community expectations that performers must accommodate while acting on some object of learning." [14]

Most children have shorter attention spans when it comes to listening to lectures and are usually more engaged in activities. Children learn from play. For this reason, the teacher we worked with when conducting our research asks the children to feed a bird puppet or to prepare a healthy lunch in her health class for pre-school children. In this way, she changes the lecture into an activity, a more engaging constructive learning environment. When children help parents or teachers to shop for groceries or cook in the kitchen, they learn about nutrition and diet through these activities. In several obesity education programs, grocery shopping and playing kitchen are used [15].

3.3 Children and Computers

Because children learn through their bodies and through active manipulation of their environments, screen-based activities at the computer are less intuitive, and thus not as effective as manipulatives, in developing understanding and skills in the early years. [16]

A kindergarten ethnographic study [17] revealed that in computer-mediated environments, objects tend to stay on the screen and cannot easily flow from computer mediated activities to other kinds of activities. The author suggests that the design of playful technology should "embody some creative and /or constructive capability and provide opportunities for discovery-oriented, autonomous play experiences."

The new generation of technology with tangible interfaces is well suited to the constructivist approach to education and thus "might encompass children in different learning spaces and promote discovery, delight, curiosity, creativity, self-expression and pleasure in learning" [18]. One good example of this kind of applications is robota, a series of doll-shaped humanoid robots, which were used as assistive technologies in behavioral studies with low-functioning children with autism. [19]

4 Design Concept and Evaluation

4.1 Design Concept

Basing on the above insights we compiled from literature reviews, interviews, and direct observations, we designed our interactive educational toy, "Smart Nancy."

Smart Nancy is a dragon character with a personality that was designed to appeal to children and to function as a role model from whom children could learn. As a long-term goal, we envisioned Smart Nancy as a cross-media character, in much the same way that Cookie Monster and Ronald McDonald are seen on TV, in books, in games, and so on.

For this study, we focused on an interactive tangible version of the character. This toy form of Smart Nancy can recognize food items and give voice feedback about whether particular food choices are healthy or not and how to make it healthier. Children interact with her by feeding her; food items include plastic representations of common food items, from bananas to cheeseburgers, which Smart Nancy is able to recognize and respond to via hidden RFID tags.

4.2 Prototype

We built an appearance functional prototype for the study. We dressed one puffy dragon we bought from the toy store and hid a speaker inside it as showed in Figure 1. When children pretended to feed simulated food items in the form of either plastic food or pictures of food, we played corresponding voice feedback from a computer.

Fig. 1. Appearance functional prototype for the study

4.3 Evaluation Study

We conducted two studies to evaluate the Smart Nancy design: one informal study and one formal study.

One 5-year-old boy participated in the informal study (Figure 2). In this study, we didn't give a health lecture or tell him Smart Nancy's story before the study and pictures of food were used instead of plastic food.

Two girls (ages 4 and 8) and two boys (ages 10 and 12) participated in the formal study (Figure 3). In this study, we gave a health lecture and told them Smart Nancy's story first, then we asked the children to feed plastic food to the dragon. Pre- and post-study surveys were conducted in order to get the information about their current eating behavior and to evaluate whether playing with the toy will influence their eating behavior.

Fig. 2. Setting for the informal study

Fig. 3. Setting for the formal study

4.4 Results

Some of the highlights of our study results indicated the following:

All the children were engaged in playing with the toy and younger children were more engaged than older ones. The 5 year old boy in the informal study tried all the pictures of food more than once. He listened to the voice feedback very carefully. He touched and hugged the dragon. All the food items in the formal study were also fed to the dragon more than once. The 4 year old girl stayed very closely to the dragon and watched it very carefully all the time.

There were a lot of interactions between the children and the toy, and among the children. The dynamic interactions in the formal study were very interesting. Figure 4 is a typical scene in the study, in which the children's eye contact reveals the extent to which they are paying attention to each other.

All the children were thoughtful in choosing the food items, and the order they fed the food items to the dragon were revealing. One child chose to feed her favorite food to the dragon first while another child chose to feed the dragon "bad food" first in the formal study. These interactions reveal how children use play to structure learning, a common pattern that the tangible toy accommodated.

Play might invoke their consciousness of healthy diet by the reflection on their current eating behavior or social pressure. Three children in the formal study happened to be siblings. We observed that one child tried to convince her sibling to choose better food in her real life during the study; for example, she encouraged her sibling to feed the dragon her favorite food. In the informal study, after he fed the dragon "grapes," he told us that he loved grapes and asked to buy more grapes for him.

Pictures of food were not as memorable as plastic food. In the informal study, the boy kept asking "Have I tried this?" In the formal study, which used plastic food, no one asked this question. This would further support the notion that tangible interfaces,

rather than 2D symbolic interfaces, are more developmentally appropriate for children.

The children played with the food items during the study. In Figure 5, the two boys made a sandwich with the plastic food.

Not all children paid attention to the whole voice feedback. At beginning of each food item, all the children paid attention to the voice feedback. However, after a while, the other children, who were not in control of the toy, began to do other things. We think that the script for each food item might be too long.

Fig. 4. Interactions during the study

Fig. 5. Two boys playing with the plastic food

5 Discussion

5.1 Refinement for the Design

After conducting and analyzing results for the study, we targeted some refinements for the next iteration of the design. These included the following:

- *Decreasing the length of the voice feedback*, which seemed to be less compelling to the participants and perhaps seemed too much like a lecture
- *Providing feedback on combined food*, to take into account playful behaviors, in which children combine different foods just to see how Smart Nancy would react.
- *Recognizing real food or scanning food labels*, so that children can more directly connect their eating habits and their consequences.

5.2 Strategy

The toy seemed to work very well, based on the interviews we did after the study. It is engaging and educational. But will it really change children's food choice in their real

life, in the short or long term? The early rounds of prototyping and evaluation cannot, of course, answer this question. We hope that once the prototype design stabilizes, we can start doing more longitudinal research to answer these questions.

As part of our marketing strategy, we think that we need to provide different concepts targeted at different age groups, cultural groups, genders. We also need to target different levels of individual factors, including social environments, physical environments and macro-systems. Certain highly popular toy lines, such as the American Girls collection, which features dolls, books, clothing, and often complex narratives, already feature these characteristics, which are presumably a source of their commercial success.

We also believe a multi-varied approach along these lines is needed to for obesity education. Our introduction of Smart Nancy as a single character is intended to compete for children's attention as a positive influence, in contrast to commercial characters from the fast food industry such as Ronald McDonald. We hope to partner with state and federal agencies to pilot a Smart Nancy campaign that would be broadly appealing.

Acknowledgements. We'd like to thank Eli Blevis, Kay Connelly, Lloyd J. Kolbe and Marty Siegel for their support of this project.

References

1. http://www.who.int/mediacentre/factsheets/fs311/en/index.html
2. http://www.cdc.gov/nccdphp/dnpa/obesity/
3. http://www.who.int/dietphysicalactivity/en/
4. At A Glance — Physical Activity and Good Nutrition: Essential Elements to Prevent Chronic Diseases and Obesity (2005)
5. http://www.who.int/dietphysicalactivity/goals/en/index.html
6. http://www.cdc.gov/healthyyouth/overweight/index.htm
7. Vandewater, E.A., Shim, M., Caplovitz, A.G.: Linking Obesity and Activity Level with Children's Television and Video Game Use. Journal of Adolescence 27, 71–85 (2004)
8. Kautiainen, S., Koivusilta, L., Lintonen, T., Virtanen, S.M., Rimpela, A.: Use of information and communication technology and prevalence of overweight and obesity among adolescents. Int J Obes (Lond). 29(8), 925–933 (2005)
9. http://www.nutritionandkids.net/1427/
10. http://www.mypyramid.gov/kids/kids_game.html#
11. http://www.hasbro.com/furby/
12. http://news.bbc.co.uk/1/hi/entertainment/tv_and_radio/4432415.stm
13. Bandura, A.: Social Learning Theory. General Learning Press, New York (1977)
14. Jonassen, D.H., Rohrer-Murphy, L.: Activity theory as a framework for designing constructivist learning environments http://www.springerlink.com/content/p32303k135654844
15. http://www.nih.gov/news/WordonHealth/jun2002/childhoodobesity.htm
16. Yelland, N.: Technology: changing the way we think and learn or maintaining the status quo. Australian Educational Computing 12(1), 3–8 (1997)

17. Wyeth, P.: Ethnography in the Kindergarten: Examining Children's Play Experiences. In: Proceedings of CHI2006 Conference on Human Factors in Computing Systems pp. 1125–1128 (2006)
18. Plowman, L., Stephen, C.: A benign addition? Research on ICT and pre-school children. Journal of Computer-Assisted Learning 19(2), 149–164 (2003)
19. Billard, A., Robins, B., Dautenhahn, K., Nadel, J.: Building Robota, a Mini-Humanoid Robot for the Rehabilitation of Children with Autism. the RESNA Assistive Technology Journal (In Press) (2006)

2D Numeric-Based Voice-Driven Browser for the Visual Impairments

Chi Nung Chu[1], Yu Ting Huang[2], and Yao-Ming Yeh[3]

[1] China University of Technology, Department of Management of Information System, No. 56, Sec. 3, Shinglung Rd., Wenshan Chiu, Taipei, Taiwan 116, R.O.C.
[2] Shih Chien University, Department of Music, No.70 Ta-Chih Street, Chung-Shan District,Taipei, Taiwan, R.O.C.
[3] National Taiwan Normal University, Graduate Institue of Information and Computer Education, 162, He-ping East Road, Section 1, Taipei 106, Taiwan

Abstract. As the aging populations are increasing, there is a need to support non-visual interaction for them in surfing the World Wide Web. This paper discusses the efficacy for the design of Voice-Driven Browser, a voice-controlled, HTML document reader/browser for the WWW. Voice-Driven Browser comprises three significant components: voice recognition, voice generation and hypertext navigation. The purpose of the integrated voice-driven browser with 2D screen reader design is to provide aging people with convenience and efficiency that will bring the WWW to those with impaired vision problems.

Keywords: 2D Numeric-Based Voice-Driven Browser, Visual Impairments.

1 Introduction

Access to the World Wide Web is one of the major problems visually impaired people are confronted with. There are about half population connected to the Internet in Taiwan [4]. The WWW have become part of our daily life. As a key tool to get into the Internet, the graphic web browser such as Microsoft Internet Explorer and Netscape Communicator is what the most of people use in Taiwan to surf the WWW. However, the aging process will affect an individual's ability to use the standard Graphical User Interface (GUI). It turns using these traditional main stream browsers visually into obstacle for them to read the information from the World Wide Web. The browser with voice which can replace the mouse in most instances to enable hands-free browsing [9].

Sighted people depend upon many visual clues to help them find information on the Web [9]. It is a complex and contextual process developed from the individual's conceptual model. There is a need, therefore, to extract the key factors of web page reading for the design of web browser with voice to provide efficient access to the visually impaired.

This paper describes a preliminary design carried out with a specially modified web browser for the visually impaired. It focuses on improving the reading efficiency of voice browsers used by the visually impaired presently. That means to transcend

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 869–873, 2007.
© Springer-Verlag Berlin Heidelberg 2007

the limitation of linear manipulation of web page reading. The idea is reduce the aural listening waiting by a two dimension model to support the users' reading strategies with voice interaction. And the existed web site can be read directly by the 2D Voice-Driven Screen Reader with no reconstruction to meet the visually impaired special needs. As the reading efficiency promoted, the author expects to get the visually impaired people up and get them running with the World Wide Web..

2 Related Works

Traditionally, web browsers are designed to interprete HTML file with the actions activated by mouse and keyboard. For the visually impaired, however, the use of alternative voice input and output might prove efficient for them to access the web sites [2]. Presently there are systems which allow voice-driven navigation, some with both voice-in and voice-out, and some allowing telephone-based web access.

The telephone-based web access uses the telephone keypad and voice command as an interface to navigate suitably configured pages. The users can then listen to the information from the handset. Phone-based browser implements the strategy to deploy a specific markup language such VoiceXML for speech rendering [6]. This approach allows the visual impaired people directly to interact with Web sites via spoken commands, and listen to prerecorded speech, music and synthetic speech.

Computer-based voice browsers, such as IBM Home Page Reader [5] and JAWS for Windows, are mainly synchronized with a moving cursor to make reading about web pages. This kind of standard HTML browser usually works in a sequential way. That means scanning from top to bottom and from left to right. Users can hear everything from descriptions of graphics to text in column format. Voice-activated browser also can navigate spoken selection of links within an extra frame with link lists extracted from the active page, such as ConversaWeb and BrookesTalk [1] [8].

However, even telephone-based web sites are built with more powerful voice-activated mechanism. It is intended for proprietary Internet Service Providers. And this new kind of markup language technology is still unfamiliar to lots of webmasters to reconstruct their web sites to meet the special needs of the visually impaired. These are the limitations to the telephoned-based web access for the visually impaired.

Intrinsically computer-based voice browser navigates web page in speech out for the visually impaired in a liner manner. The main distress for the visually impaired comes from waiting the voice browser to get the right place to read out the desired information. That would be a barrier for the visually impaired to access the web sites. Furthermore the voice browsers are mostly designed in English version. That is also a constraint for the non English users to manipulate.

3 2D Numeric-Based Voice-Driven Browser

The design of 2D Voice-Driven Browser (Fig. 1) is based on the Microsoft Internet Explorer engine running on the Windows platform. The IBM ViaVoice Chinese text-to-speech synthesizer is embedded in charge of speaking web-based information aloud. 2D Numeric-Based Voice-Driven Browser is integrated into standard graphical

browser and installed at the client side. A range of adaptive interface, such as tool bar position arrangement, voice description for each function button and magnifier, is accommodated at the interface [3].

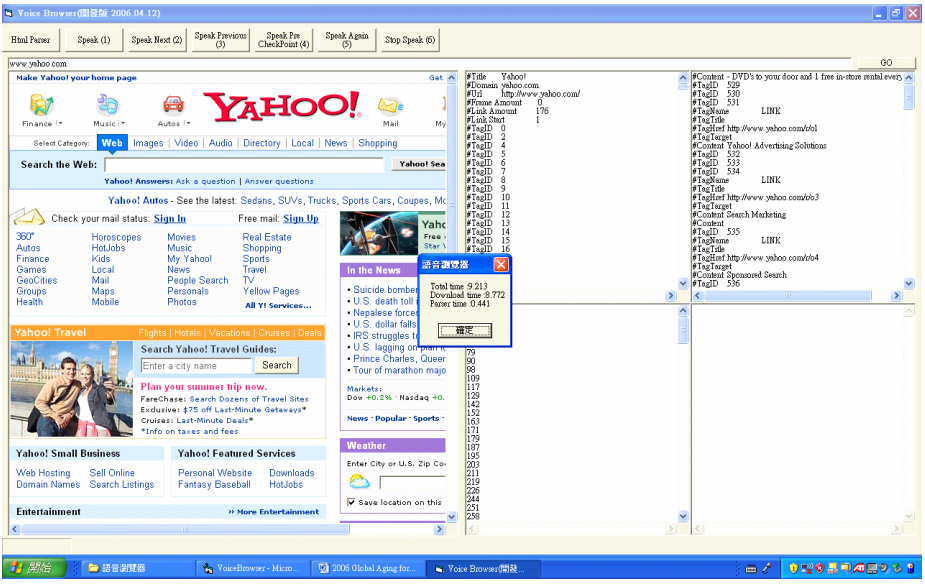


Fig. 1. 2D Numeric-Based Voice-Driven Browser

3.1 Voice Interaction

Current web page structure consists of frame, table and form with contents including hyperlinks in each part. In order to provide better web page information access for the visually impaired, 2D Numeric-Based Voice-Driven Browser will first speak out the structured parts of a webpage such as frames, tables or forms with their corresponding titles. In the meantime, user can either choose 8 (speak or press) to listen to further contents in current readout part or continue listening to the next parts of the webpage. The number-based trigger mechanism is as follows:

0 URL input
1 skipping for next object table, frame, form
2 skipping for previous object
3 previous check point
4 speaking current object repeatedly
5 pause
6 previous page
7 next page
8 selecting current speaking object (including hyperlinks)
9 next check point

With the above mechanism, the target area of a web page can be focused just like what the sighted people do when they navigate the web page.

3.2 2D Vocal Web Page Content Creations

2D Numeric-Based Voice-Driven Browser reads out web pages to users via Chinese Text-To-Speech engine. The visually impaired could control the 2D Numeric-Based Voice-Driven Browser navigation by choosing a number just like listening to radio. 2D Numeric-Based Voice-Driven Browser can analyze HTML files in accordance with the W3C WAI Guidelines [7]. According to these simple guidelines, the website masters can smooth the voice browsing perception for the voice voyagers.

The 2D structure is rendered by analyzing some specific HTML tags and attributes within the tags. In order to make a voice triggered HTML page friendly for the visually impaired to navigate, more clear tags and attributes could produce better descriptions by the 2D Numeric-Based Voice-Driven Browser to generate readout dialogue. Some key tags and attributes are interpreted by the 2D Numeric-Based Voice-Driven Browser as the following:

1. <TITLE>: The TITLE tag displays the title of an HTML document on the browser's title bar. 2D Numeric-Based Voice-Driven Browser speaks it out first for users as it downloads one page from a web site.
2. Attribute *summary* in <Table>: This attribute provides a summary of the table's purpose and structure for 2D Numeric-Based Voice-Driven Browser rendering aural description to the users.
3. Attribute *title* in <Frame>: 2D Numeric-Based Voice-Driven Browser reads the frame title to facilitate frame identification and navigation for the users.
4. Attribute *alt* in <IMG>: 2D Numeric-Based Voice-Driven Browser reads out image representation of text to the users.

The use of attribute *alt* is often ignored by web site masters since it is unseen on the foreground. However, that is extremely important information for the voice browser to speak out for the visually impaired. There are still many other such attributes available in the HTML tags which can be found in the W3C WAI Guidelines for voice browser to interpret.

4 Conclusions

The goal of this design is to find an efficient solution for the visually impaired to surf the World Wide Web. 2D Numeric-Based Voice-Driven Browser implements voice interaction technology to make a wealth of information available on the Internet for the visually impaired by choosing only one of number 0 to 9. This paper discusses the need and design of 2D Numeric-Based Voice-Driven Browser -- a voice-controlled document reader/browser for the World Wide Web to support those with sight disabilities. It is great for the visually impaired to have the opportunity reading the web pages as the sighted people do when they navigate the World Wide Web. Spoken number facilitates voice recognition. With this approach the web designers have not

to reconstruct web pages on the part of server and the visually impaired could listen to the web page efficiently with their own mother tongue.

References

1. Adaptive Technology Center Indiana University Bloomington, http://www.indiana.edu/ iuadapts/ convers2.html
2. Hemphill, C., Thrift, P.: Surfing the Web by Voice. In: Proceedings of ACM Multimedia, San Francisco, CA, November 7-9, 1995 pp. 215–222 (1995)
3. Chu, C.N., Chen, M.C., Li, T.Y.: The Design of an Adaptive Web Browser for Young Children with Reading Difficulties. In: Miesenberger, K., Klaus, J., Zagler, W. (eds.) ICCHP 2002. LNCS, vol. 2398, pp. 189–190. Springer, Heidelberg (2002)
4. Computer Industry Almanac Inc. USA tops 160M Internet Users (December 16, 2002) http://www.c-i-a.com/pr1202.htm
5. IBM Accessibility Center http://www-306.ibm.com/able/
6. Voice Extensible Markup Language (VoiceXML) Version 2.0. http://www.w3.org/TR/voicexml20/
7. Web Content Accessibility Guidelines 1.0, W3C http://www.w3.org/TR/WAI-WEBCONTENT/
8. Speech Project Oxford Brookes University http://www.brookes.ac.uk/schools/cms/research/speech/ btalk.htm
9. Zajicek, M., Powell, C., Reeves, C.: A Web Navigation Tool for the Blind. In: Proc. 3rd ACM/SIGAPH on Assistive Technologies, Los Angeles, USA (1998)

Case Study of Applying Wireless Technologies into Healthcare Industry in China and UK

Guixia Kang[1], Li Zhang[2], Shanghong Li[1], Ping Zhang[1], and Said Boussakta[2]

[1] Wireless Technology Innovation (WTI) Institute
Beijing Univ. of Posts and Telecommunications
[2] Institute of Integrated Information Systems (I3S)
School of Electronics and Electrical Engineering
University of Leeds
gxkang@bupt.edu.cn

Abstract. In recent years, there have been growing interest and activities of introducing wireless technologies into healthcare industry. Successful development and deployment of wireless e-Health (WEHealth) systems is geographically dependent in respect of national regulations, regional economic development, and available medial infrastructures and resources, etc. In this paper, the applications of wireless technologies into healthcare industry are investigated in China and UK respectively. Based on the case study of wireless healthcare in China and UK, this paper will analyze the impact of above mentioned factors on the integration of wireless technologies and healthcare applications, and propose possible solutions based on the analysis.[1]

Keywords: e-Health, wireless, healthcare, homecare.

1 Introduction

In recent years, there have been growing research interest and activities of introducing wireless technologies into healthcare industry. On one hand, the demand of improving the efficiency and diversity of healthcare services is increasing. Applying the wireless implanted medical sensors enables the monitoring and control of the medical conditions such as coronary care, diabetes, muscle stimulants, etc. [1]. The increasing demand to measure physiologic responses in patients' normal activity status (such as rest, sleep and slow walking) for more realistic indications of the patients' health requires the use of wearable wireless monitoring devices which allow continuous or intermittent monitoring of physiologic signals [2][3]. On the other hand, the fast development of wireless technologies, especially the wireless sensor and networking techniques, has given sensors various functionalities [4], which allow the sensors to be the end node in the ubiquitous computing environment. The fast development of wireless technologies (such as Bluetooth, Zigbee, UWB, and WLAN) makes it more

[1] This paper is sponsored by Beijing Nature Science Foundation (No. 4062023) of China, the Royal Society 2005/R3 of UK, and the 111 Project （No. B07005） of the Ministry of Education (MOE) of China.

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 874–882, 2007.
© Springer-Verlag Berlin Heidelberg 2007

and more attractive and suitable for medical applications [5]. Soon after the appearance of a multi-disciplinary concept of e-Health in Europe North America and Asia, the wireless applications into e-Health industry are widely investigated [6]. This enables the wireless professional real-time patient monitoring and therapy support as well as private user access to monitoring, medical assistance and individual care services in future pervasive/ubiquitous computing environments.

In this paper, the applications of wireless technologies into healthcare industry are investigated in China and UK respectively. It is known that the aging population, health situation, social services and medical tradition are different in western and eastern countries, and these all affect the development of wireless e-Health. It has been recognized that successful development and deployment of the e-Health systems is geographically dependent in respect of national regulations, regional economic development, culture and available medial infrastructures and resources. This paper will analyze the impact of these factors on the integration of wireless technologies and healthcare application, and propose possible solutions based on the analysis results.

This paper is structured as follows. In Section II, the characteristics of healthcare in China and UK are overviewed. In Section III, the preconditions of successful applications of wireless e-Health (WEHealth) in China and UK are discussed. In Section IV, based on above discussions, the possible application scenarios of WEHealth in China and UK are investigated. In Section V, the exemplary WEHealth solutions in China and UK are introduced. In the last section, a joint conclusion will be drawn based on the above study.

2 Characteristics of Healthcare in China and UK

Successful development and deployment of wireless healthcare systems are geographically dependent on national regulations, regional economic development, and available medial infrastructures and resources, etc. There are always increasing needs of healthcare in terms of the healthcare facilities, professionals, workers, and services. China has the largest healthcare systems in the world including all sorts of medical institutions of 298,997 by the end of 2005, in which there are 18,703 hospitals, 17,129 community healthcare service centers, and 41,694 country healthcare sanatoriums, as well as 5,426,851 professional workers on healthcare [7]. However, comparing with the whole population, which reaches 1.29 billion by the end of 2005, the healthcare resource is still quite scarce. When considering the shortage of resources of hospitals relative to the whole population, the healthcare in China is encountering reforms to move traditional hospital care to community care and even homecare. When considering the coming aging society and China's national policy of One-couple One-child, the healthcare on the elderly population will become a severe problem for Chinese society, although currently the elderly population is 10.97% of the total population by the end of 2004 [8], which is quite below that percentage in European countries. For example, in UK, the aging rate is 16.7% up to mid-2005 [9]. The aging issue is a common social problem to both China and UK. In addition, in the wide relatively poor area such as in western China, the healthcare coverage is far below the average level. The healthcare improvement in this area

should particularly consider the available resources and facilities. An important fact to this area is that the accessibility to PC for the families ranks below that of mobile phones. For example, in the urban area of western China, only 42% people have used or heard of computer or Internet; while the number for mobile phone is more than 80%. Even in the rural area of western China, around 80% people have used or heard of mobile phone, comparing to merely 24% for computer or Internet [10]. Actually until September 2006, the mobile phone users in China has exceeded 0.443 billion. In the developed European country such as UK, more than 80% of the population uses a mobile phone. The popularity of mobile phones in both China and UK makes it a promising platform for healthcare services. However, when we consider the suitable facility to spread healthcare services in both the urban and rural area, especially in hospital applications, we have to consider variable factors, such as the regulations of the hospitals. For example, in the UK, the use of mobile phones is limited or banned in hospitals. In China, the unbalanced healthcare and service level between the large highest level (AAA level) hospitals and community healthcare service centers, the unbalanced healthcare facilities between urban and rural area, the nomadic labors entering into the cities, etc., should be considered. In the UK, the National Health Service (NHS) was set up in 1948 and is now the largest organization in Europe. It is recognized as one of the best health services in the world by the World Health Organization. However, there need to be improvements to cope with the healthcare demands of the 21st century. The NHS now has many problems, such as the seriously shortage of staff and resources, and the long waiting list to treatment. In summary, when we seek for healthcare solutions in different countries, especially when we hope to have some general solution to different countries, some key characteristics of them have to be taken into account.

3 Preconditions of Wireless e-Health in China and UK

Although WEHealth solutions have many benefits such as flexibility and wide coverage as well as accessibility, some preconditions should be satisfied in China for a successful deployment of WEHealth systems.

1) Well-developed wireless e-Health products provided by the manufactures
Although there might be lots of deployment possibilities of WEHealth systems, from the end-users point of view, they will expect the WEHealth systems to be cost-effective, safe, wearable, energy-effective, and privacy protected. The terminals can vary from the simple configuration for transmission of low data rate, e.g., medical data such as blood pressure, heart-rate, etc., to relatively complicated configuration for transmission of medical pictures. The end users of WEHealth products include the patients, the doctors and nurses, and some third party of service providers in China, probably the community workers and professionals.

2) Public recognition by hospitals, communities, end-users (e.g. elderly people at home), and mobile operators
In Sept. 2005, sponsored by the China 863 program, the 1st domestic conference on e-Health in China - Beijing Conference on e-Health, was successfully held in Beijing. The attendees to this conference came from the government such as the Ministry of

Science and Technology of China (MOST), the Ministry of Healthcare (MOH), the National Population and Family Planning Commission (NPFPC), etc., from the research institutes, such as Beijing University of Posts and Telecommunications, Tsinghua University, etc., from the hospitals, such as TongRen Hospital, Center Hospital of Guang Zhou Military Area, ChaoYang Hospital, etc., as well as from the Chinese companies of medical devices. The attendees widely discussed the possibility of e-Health in China and believed that it will be realized in the next 5-years. Wireless technologies are recognized as one of the promising technologies that can be used together with e-Health. In future, in order to obtain wider public reorganization of wireless e-Health, information exchange between research institutions, hospitals, end users (e.g. elderly people at home), and maybe mobile operators should be promoted.

3) Government promotion

The wide deployment of wireless e-Health systems needs the promotion from the government by acknowledging some key benefits by wireless e-Health including:

Medical Cost Reduction: the rising medical cost has become a big problem not only in China but in many other countries such as in Japan, Switzerland and the UK. In these high expenditure countries, many people lose confidence to the hospital. Wireless technologies can help to reduce the medical costs by, e.g., more convenient information access for the patients.

Mitigating Aging Society Pressure: by introducing wireless transmissions together with sensor techniques, the elderly people can be monitored even at home. In this way, a lot of costs can be saved compared with they might spend on the nurses.

4) Multi-disciplinary and international cooperation

The development of the WEHealth prototypes and products calls for the multi-disciplinary and even international cooperation. It is recognized that the ICT experts are not the only key personnel in developing the WEHealth products. The end users, including the patients, the community workers and the doctors/nurses should take part in the R&D of WEHealth systems from the initial stage. The requirements and concerns from the end users will be the deterministic factor for a successful WEHealth system. Sometimes the doctors and nurses have more conservative expectations than what the ICT experts have imagined.

Besides, the international cooperation is important to learn the success or failure experienced in other countries.

5) Step by step deployment

From city to rural area: To realize the wireless e-Health system throughout China is a huge project. Generally the city deployment might probably priories the deployment in rural area. However, there might be some extra case that the rural area has the deployed system simultaneous with urban area. For example, NPFPC is promoting the wireless transmission of the population and health data in the wide rural area, since the population statistics such as the birthrate is more important in rural area than in the city.

Upgrade existing system gradually: Doctors have the habit to do something they are familiar with. If the new system completely changes existing systems, doctors won't use it at all even though the new system may have many benefits. So the new system must be based on existing system and upgraded gradually.

A nation's health service is fashioned by its economy, demography, culture, and medical tradition. Hence, in the UK, the preconditions for wireless e-Health applications are slightly different, although they may have similar headings.

1) Well-developed wireless e-Health products provided by the manufactures

Although the wireless technology has been well developed, to apply to healthcare industry, a flexible and integrated telemedicine linkage, and compatibility between the telecommunication services and the current standards are required. For example, the high cost of communication links, the limited data transfer rate of the current mobile phone systems and the limited availability of mobile internet connectivity and information access will hinder the application of wireless technology in healthcare industry. The increasing data traffic and demands from different medical applications and roaming application must be compatible with the data rates of wireless systems in specific mobility conditions. There is also lack of integration between wireless systems and other information systems used in medical systems, such as referral and ordering systems, medical records etc. Therefore, to promote WEHealth, the well-developed wireless products suitable for medical applications must be readily available.

2) Organizational changes

Healthcare is a very complex industry and difficult to change. It is well known that current Healthcare systems are stuck with the equation:

Current organization + new technology = expensive current organization.

Therefore, organizational changes are very often required for healthcare institutions to benefit from Wireless technology. Those required organizational changes most likely have an impact on how physicians and other staff members lose or gain power as a result of those changes. Meanwhile, the short term and long-term economic consequences and working conditions for physicians and healthcare systems must be fully understood and taken into account. In addition, the methods of payment for such WEHealth system should be fully developed and standardized.

3) Recognition by hospitals and end users

The new technology must be recognized by hospitals and end users before their deployment. However, there are not enough numbers of demonstration projects that show WEHealth's real potential.

4) Step by step deployment

As aforementioned, the organization change is essential for the deployment of WEHealth, therefore, the existing healthcare system should be changed gradually. Another reason for this is the lack of incentive for busy specialists to practice the integrated new wireless technology: the currently available WEHealth equipment can sometimes be difficult to handle, because it is seen as yet another imposition for which they are not paid.

4 Wireless e-Health Scenarios in China and UK

Our vision of the wireless applications in healthcare industry in China and UK reflects the above analysis and concerns by considering the characteristics of healthcare and

preconditions in these two countries. Basically we envision the following scenarios for wireless e-Health development and deployment in these two countries:

4.1 Home Monitoring

Home healthcare is a rapidly growing component of today's health care system. More services can now be delivered in the patient's home, but professional supervision and monitoring of vital functions and medications are still required, rural and less served urban patients also face travel and access problems. Our vision on WEHealth homecare strengthens the applications of short range and low power wearable wireless sensor networks to monitor the physical parameters of the patient, e.g. blood pressure, blood oxygen saturation, body temperature and/or ECG on a relatively long-term basis for the aging people with chronic diseases, such as asthma, cardiovascular, and diabetes. Zigbee can be chosen as the technique for transmitting the patient's physical data to the remote monitoring center in hospital, since even for the ECG transmission, a data rate of tens kbit/s should be sufficient for the graph transmission.

Fig.1. illustrates a Home Monitoring scenario. To make the home monitoring efficiently run, the community medicine center and even hospital are involved. Based on the home monitoring (on one or a few patients), the application environment can be extended to the whole community. The daily health monitoring can be carried out at the community, while the professional diagnosis should be more possibly provided by the hospital. The patient's data come from sensors or automation carried by person or set on certain place at home, e.g., embedded in bed, chair, and shoes. Wireless sensors collect the people's vital signs (blood pressure, heart rate, etc), and then relay to family sink/gateway.

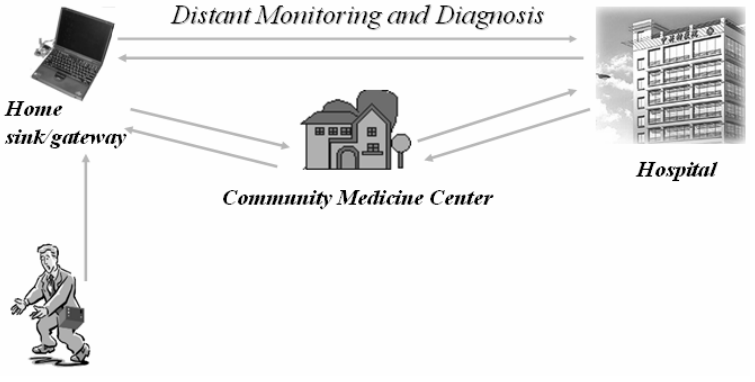


Fig. 1. Home Monitoring Scenario

Since both China and UK are encountering the aging problem, and the homecare with special nurserymaid is expensive (more expensive in UK than in China); and because the regulation of wireless applications at home generally is less strict than at hospital, the WEHealth system can be more necessarily and easily applied as home monitoring.

4.2 Community Scenario

In the community monitoring range, the patient can go out of home to do some exercise, buy food or chat with friends. In all these activities, the health status can be continually monitored. All the real-time data can be transmitted to the community monitoring center and analyzed. Then the data can be transferred to hospital's patient medical record database. If something abnormal happens or the monitor finds the value trend abnormal, the monitoring server in the community will alarm patient and notice hospital request instruction. Besides the health monitoring in the community range, another important content of community care is the disease prevention based on the broadcasting of the epidemic information which can be carried on mobile phones. Some other WEHealth services can be provided in the community area.

Since in recent years the public sanitation is paid more and more attention by the Chinese government, and because the community clinics can greatly lessen the pressure of healthcare of the hospitals especially in big cities, we can envision more and more community applications of WEHealth systems.

4.3 Hospital Scenario

WEHealth systems and services can be applied to hospital scenario as well. In this scenario, not only the traditional hospital services can be carried on mobile terminals (such as mobile phones, PDAs, tablet PC, etc), resulting in wireless electronic health record (EHR), mobile patient examination, wireless access to database, etc, but also some new functions such as tele-medicine and tele-registration can be provided. Our vision on WEHealth in hospital care encourages the application of WLAN in hospitals to simplify the health data gathering and accessing by the doctors and nurses with high data rate. With GPRS or UMTS coverage to hospital, the tele-registration and tele-medicine can be conveniently provided also.

However, as stated in the last section, the standard of the information exchange and the regulation of wireless applications in hospitals should be considered seriously if WEHealth systems are applied to hospital scenario.

5 Exemplary WEHealth Solutions in China and UK

Based on above analysis, we present some WEHealth solutions in China and UK, which we think might be applied earlier than the other WEHealth solutions.

When considering the fact that the available and familiar facilities for Chinese people, both in urban and in rural areas, are mobile phones rather than computers, some innovative and ambitious development and deployment of healthcare systems and services are promising by introducing wireless technologies into China's healthcare industry. An example of WEHealth solution in China is a WAP based remote registration system [11]. Compared with the huge number of Chinese population, the registration to the hospitals becomes a serious problem especially in some famous AAA hospitals such as in Tongren hospital, Tiantan hospital, etc. People have to queue for a long time to register; while some registration ticket packman always stand in the front of the queue and buy a lot of registration numbers

and then sell with much higher price to the patients than the normal price. This problem can be efficiently solved by introducing wireless services to the hospitals and patients. The implementation structure is shown below [11]:

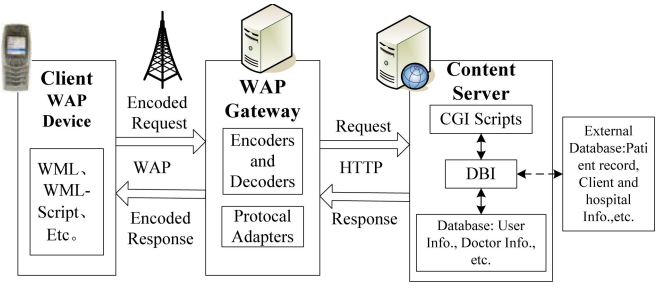


Fig. 2. Implementation Architecture of Wireless Registration System

The system basically implements the function of WAP registration by the patients and the registration management by the hospitals. The system utilizes WAP technology to realize the connection between the wireless domain and the WWW domain. The related application data were stored or generated by the content server. The human-machine interface is implemented on mobile phones with several simple steps to register to the proper doctor at the proper time. The management at the hospital server can be integrated with the current health information system (HIS) to give more information to the end users.

In the UK, the fixed line operator BT is deeply involved in the NHS connecting for health programmes. The introducing of wireless technologies can provide services such as SMS patient reminders, mobile calls to NHS direct and monitor patients in their own home etc, which can significantly increase the efficiency in NHS, reduce the demands in staff and resources, and shorten the waiting list. The issue of staff and resource shortage and the long waiting list to treatment can be mitigated by introducing WEHealth systems and services to the healthcare industry. Another solution is an obesity tackle system based on wireless technologies. It is well-known that obesity is responsible for more than 9,000 premature deaths per year in England. Obesity is an important risk factor for a number of chronic diseases such as heart disease, stroke, some cancers, and type 2 diabetes. The prevalence of obesity in children aged under 11 increased from 9.9 percent in 1995 to 13.7 percent in 2003. In July 2004 a target specifically on obesity was set for the first time - 'halting the year-on-year rise in obesity among children aged under 11 by 2010 in the context of a broader strategy to tackle obesity in the population as a whole'. A mobile based service, called MyFoodPhone, to help people adhere to dietary regimes, is started recently by NHS direct in the UK. This service will provide guidance and information for people suffering with weight problems. It is a pretty simple service, which works with following steps:

- With a camera phone, you take pictures of what you eat.
- You upload the pictures to your personal dietitian.
- Once a week, your dietitian analyzes your information.

- She then sends you a video clip of feedback.
- Follow your dietitian's advice and you should begin to see results.

This service aims at obese and is welcomed by many people.

6 Conclusion

WEHealth development and deployment is geographically dependent in respect of national regulations, regional economic development, culture and available medial infrastructures and resources. The common vision of WEHealth can be drawn in both countries such as its applications in home scenario, community scenario as well as in hospital scenario. All scenarios show that wireless technologies are promising to provide more services to current healthcare industry which can solve some of the social problems. However, it has also shown that, due to the different regulations, economic development, culture and medical traditions in China and the UK, to promote WEHealth in these countries, they have to start from different initiating solutions reflecting their characteristics of healthcare. In the future, to develop and deploy WEHealth successfully, the above factors have to be taken into account in order to find the best suitable solutions for different countries.

References

1. Timmons, N.F., Scanlon, W.G.: Analysis of the Performance of IEEE 802.15.4 for Medical Sensor Body Area Networking. IEEE Sensor and Ad Hoc Communications and Networks (SECON) 2004, pp. 16–24 (October 2004)
2. Binkley, P.F.: Predicting the Potential of Wearable Technology. IEEE Engineering in Medicine and Biology Magazine, pp. 23–27 (May/June 2003)
3. Bonato, P.: Wearable Sensors/Systems and Their Impact on Biomedical Engineering. IEEE Engineering in Medicine and Biology Magazine, pp. 18–20 (May/June 2003)
4. Sokwoo, R., Yang, B., Chang, K., Asada, H.H.: The Ring Sensor: a New Ambulatory Wearable Sensor for Twenty-four Patient Monitoring. In: IEEE Conference on Engineering in Medicine and Biology Society, pp. 1906–1909 (October 1998)
5. Ng, H.S., Sim, M.L., Tan, C.M., Wong, C.C.: Wireless Technologies for Telemedicine. BT Technology Journal, vol. 24(2) (April 2006)
6. EU FP7 Call on Information and Communication Technologies, Challenge 5 – toward sustainable and personalised healthcare
7. Communique of Chinese Healthcare Development in 2005 http:// www.moh.gov.cn/open/ uploadfile/ 200642584823110.doc
8. Research Report of the Tendency Prediction of Chinese Aging Development http:// www.china.com.cn/chinese/news/1134589.htm
9. http://www.statistics.gov.uk/CCI/nugget.asp?ID=6
10. Zhang, K., et al.: The investigation of people's demand and acceptability on eHealth in urban and rural areas in Western China. Report in the First Chinese Forum on e-Health, Beijing, China (September 2005)
11. Ouyang, Y., Kang, G., et al.: Investigation and implementation on the advanced wireless medical registration solution in China, HCII 2007 (accepted for publication) (2007)

Attitude Toward Wireless Applications in Healthcare Industry in China

Guixia Kang[1], Yue Ouyang[1], Da Liu[2], Huaqing Wang[1], and Ping Zhang[1]

[1] Wireless Technology Innovation (WTI) Institute
Beijing University of Posts and Telecommunications, P.R.China
[2] Intelligent Technology and Machine Lab
Beijing University of Aeronautics and Astronautics, P.R.China
gxkang@bupt.edu.cn

Abstract. A validated questionnaire evaluating the attitude of Chinese people toward wireless applications in healthcare industry was administered to around 1,000 people from Oct. 2006 to Jan. 2007. Investigated subjects include people with different occupations and in different age levels. Investigated content covers preferred service types of wireless applications in healthcare, people's concerns towards the wireless services, etc. Investigated methods include interviewing and dispatching questionnaires. The objective of the survey is to assess people's acceptability and requirement toward wireless applications in healthcare industry, called Wireless e-Health or "WEHealth" for short. Result reveals that people's attitude toward WEHealth in China is positive basically; An overall analysis of questionnaires reveals some interesting results and some problems we should be pay more attention to in the future industrialization process of WEHealth systems and services. [1]

Keywords: Questionnaire, wireless, healthcare.

1 Introduction

WEHealth (Wireless e-Health) [1] is an emerging field in the intersection of medical informatics, public healthcare and business, referring to healthcare services and information delivered or enhanced through wireless (GPRS, WLAN, etc) and wired technologies. In a broader sense, this term characterizes not only technical development, but also a state-of-mind, a way of thinking, an attitude, and a commitment for networked, global thinking, to improve healthcare locally, regionally, and worldwide by using advanced information and communication technologies.

Now more and more countries pay attention to their medicine treatment systems. Every country and its people expect to improve their quality of healthcare and the management and operation in hospital. Meanwhile, some countries start to consider reducing the likelihood of costly errors by using mobile technologies. In China, the aging society is coming and the aging issue becomes severer and severer due to the

[1] This paper is supported by Beijing Nature Science Foundation (No. 4062023) and The 111 Project (No. B07005) of Ministry of Education (MOE) of China.

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 883–892, 2007.
© Springer-Verlag Berlin Heidelberg 2007

national policy of One-Couple One-Child. The child is growing up to a youth with higher pressure of life than their parents and with little leisure time for the care-taking of their parents and grandparents. Therefore, a system which can lessen the youths' pressure on care-taking of their family members is solicitous. In addition, the lack of hospitals relative to the whole population and the non-ideal medical coverage in the wide rural area of China oblige us to seek efficient ways to resolve the contradictory between the increasing medical requirement of Chinese people and the insufficient medical resources. WEHealth is therefore proposed as a pioneer of the wireless applications in healthcare industry. Actually WEHealth can be interpreted into a series of systems, including remote and long-term patient monitoring via wearable devices and wireless sensors [2], WAP-based health information systems (such as remote registration system [3]), and some other systems and services such as emergency telemedicine [4]. In recent years, national policies of China have initiated to support the advanced Information and Communication Technologies (ICT) including wireless technologies to solve the social problems. It is clearly declared that Modern Service Industry, in which healthcare is one of the topics and in which wireless technologies might play some roles, is an important development domain in the guideline of National Metaphase and Long-term Developing Plan in science and technology [5][6].

Undoubtedly, people's acceptability and requirement are the most important impetus to new technologies or to the application of traditional technologies to new industry areas. To find out people's attitude toward WEHealth can provide validated evidence of how to promote the industrialization and related research of wireless technologies into healthcare industry. To investigate people's subjective desire and the perceived benefit of e-Health, this survey was administered in terms of the acceptability and requirement analysis.

This paper is organized as follows. In Section 2, the infrastructure of WEHealth is introduced. In Section 3, the method of the investigation on people's attitude toward WEHealth systems and services is introduced. In Section 4, the result of the investigation is presented and analyzed. In Section 5, some discussions and the conclusion are made.

2 Infrastructure of WEHealth

Fig. 1 gives the infrastructure of WEHealth. The term 'wireless' is strengthened in WEHealth concept since it is observed the various wireless access techniques and networks, such as short range wireless communications, wireless local area network, as well as wide range wireless communications, can be applied to provide advanced and convenient healthcare services. In the home environment, for example, the physical sensors (such as blood pressure sensor, heart rate sensor, etc) or nonphysical sensors (such as the action detection sensor) are attached on the human body to collect the parameters of the patient-in-care, typically elderly people in future aging society. The collected data are transmitted, e.g., through short range wireless (Bluetooth, Zigbee, or UWB) transmitter to the home gateway, which might be embedded into home ADSL, PC, mobile phone, PDA, etc. Through home gateway, the patient's daily data are sent to the hospital to implement the real-time monitoring

on the patient. Variety of healthcare services can be developed through the platform of the home environment [7]. A typical service is the monitoring of the chronic diseases (diabetes, heart diseases, etc). In the hospital environment, the wireless local area network can be established to improve the medical information access and exchange capability between the patients and doctors, or between the doctors and the hospital database. The electric prescription and body check result are within the abovementioned information, which can be conveniently transmitted through wireless access such as WLAN. Variety of applications can be developed in hospitals for this information access and exchange [8], and the terminals are normally PDAs or other mobile terminals with bigger screen due the amount and diversity of medical information. In office environment, the ordinary people can be the service-taker through the wide range wireless network, such as GPRS, 3G (WCDMA or CDMA 1X), etc. The services can be also diverse, such as the daily health information broadcasting through SMS, the mobile registration to the hospital [3], etc.

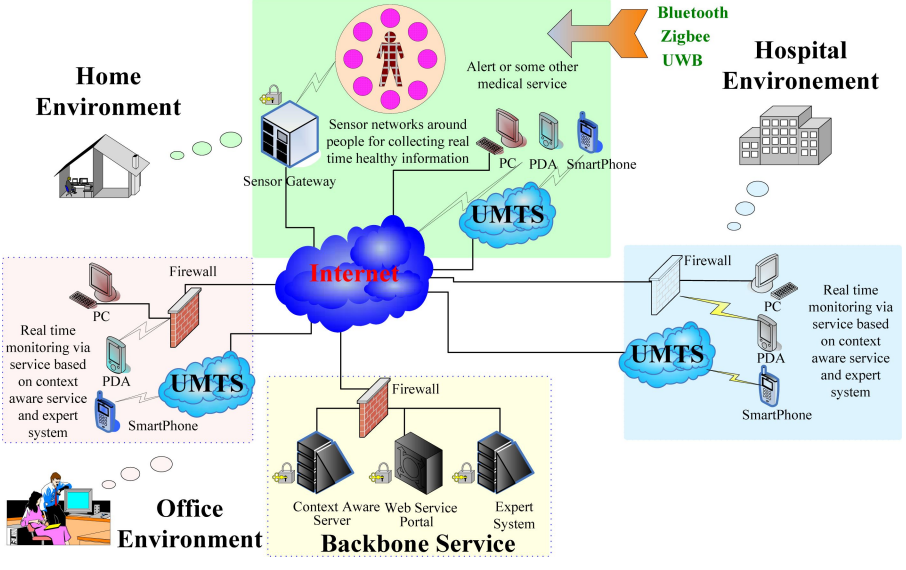


Fig. 1. Infrastructure of WEHealth

3 Investigation

The implementation of WEHealth systems might be diverse. Although basically WEHealth is to implement healthcare services based on some matured wireless technologies, these kinds of application domain is still relatively novel, and people's attitude towards WEHealth systems and services will play important roles in their successful application and deployment. In this paper, a validated questionnaire measuring the attitude of Chinese People toward WEHealth was administered to around 1,000 people from Oct. 2006 to Jan. 2007. The questions are carefully

designed to try to get some systematic and meaningful results concerning the attitude of Chinese people toward the acceptability and requirement of WEHealth in future.

3.1 Design of Questionnaire

To get further understanding on the requirement and acceptability of people, we shape the related investigate frameworks combining with the characteristics of WEHealth. Then we provide multi-profile acceptability and requirement frameworks as shown in Fig.2 and Fig.3.

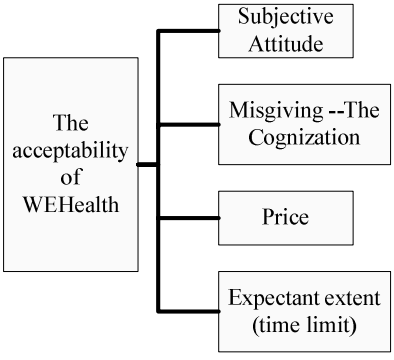


Fig. 2. Research framework of WEHealth acceptability

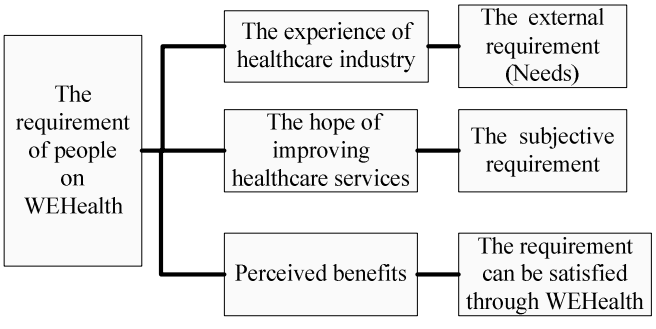


Fig. 3. Research framework of people's requirement

Exemplary questions toward the acceptability of people on WeHealth include: "Do you have the experience of accessing Internet through mobile phones?"; "When do you expect the wireless healthcare systems/services to be implemented?", etc. Exemplary questions toward the requirement of people on WeHealth include "Which kinds of functions do you expect the monitoring systems/services to provide?", "which kinds of services do you expect if you can access Internet through mobile phones for healthcare?", etc.

3.2 Method of Investigation

We use accidental or convenience sampling method [9] by questionnaire and interviewing. Accidental sampling is also known as incidential sampling, haphazard sampling, etc. In this sampling technique the researcher interviews all the respondents for that study who comes in contact accidentally with during the research time. In convenience sampling, the selection of units from the population is based on easy availability and/or accessibility. Questionnaires were delivered to subjects in university, hospital, park, community, and company. The content in questionnaire includes:

a) people's experience of related healthcare services;
b) people's perceive of benefits from WEHealth applications;
c) evaluation on the acceptability of WEHealth and on the attitude toward WEHealth applications, and the perceive of WEHealth's characteristics;
d) the external requirement of people in order to improve current healthcare services.

To reduce sensitivity of questionnaire, we promise using anonymous method and will not reveal the personal information.

Selecting the survey subjects is an important element of any research of this type. The target group includes three types: students, employees and aging people:

a) student: although they are studying in universities and generally do not have any income, they might probably become the potential customers of WEHealth services in the near future, especially for some advanced health information systems based on wireless application protocol (WAP) services;
b) employees: they have steady income, working under high social pressure; they have parents and grandparents who are stepping into elder ages who have retired (or will retire soon); they currently are the main consumer group of the potential WEHealth systems since they need to take care of their parents, themselves as well as their children;
c) aging people: most of aging people suffer from various chronic diseases; The chronic disease might not threat their life immediately. But the physiological parameters of the aging people need to be monitored and controlled since some emergent attack of the disease might happen. These persons will be the fundamental and targeted users of WEHealth systems. This group of people will provide their important experience on traditional healthcare services and their attitude toward WEHealth services.

Consequently, the characteristic of the questionnaire is that we start with people's experience on healthcare related services and then on the basis of introducing the concept of WEHealth, we ask people to expect the realizable healthcare services, express their wishes, judge and master the accessible benefits.

4 Result Analysis

We get a deep understanding on people's subjective opinion of WEHealth by questionnaires and scene investigation. As mentioned above, this survey takes the

accidental or convenience sampling method in Beijing. The age of investigated people is mostly over 20 years old. The average education degree is relatively high. Bachelor degree takes the percentage of 80%, master degree takes 15% and doctor degree takes 4%. The gender rate of male takes percentage of 63.9% and female 36.1%. The analysis of statistic data is given in the following.

4.1 Requirement Analysis

Fig.4 shows that chronic diseases suffer most Chinese families; the chronic diseases' sufferers take the percentage of 67%, in particular, hypertension 32%, diabetes 15% and cardiopathy 13%. So in the aging society of China, the trouble of chronic diseases has already been a main problem to most families, and how to take care of these chronic diseases' sufferers is going to be a prevalent social problem. Whether this problem will be solved, as far as personal entity is concerned, it will influence people's life quality directly and as far as social colony are concerned, it will influence harmonic social development directly. According to Chinese economic development level nowadays, most families can not afford to hire special paramedic to take care of these family chronic diseases' sufferers. In addition, rarely families have or intend to hire special person such as nurserymaid to take a long-term care on patients. These situations are confirmed in Fig.5.

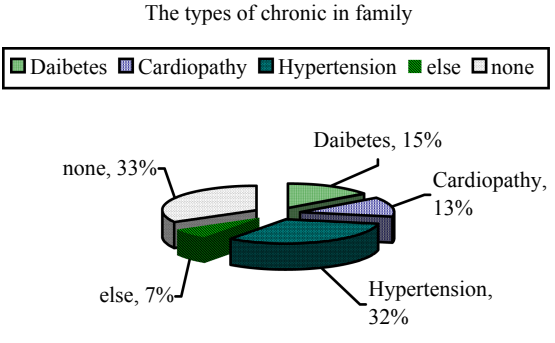


Fig. 4. Types of chronic diseases in family

In general, the wireless monitoring for chronic diseases' sufferers and their families has a huge market potential, which is the external strong impetus of WEHealth systems and services. After the explanation of WEHealth and related wireless applications to the investigated subjects, they express their attitude toward the necessary of the monitoring service and other WEHealth services. The result on people's attitude on the monitoring service is shown in Fig.6. We can see that 89% investigated persons consider they will need the monitoring service or equipment for family chronic diseases' sufferers, only 11% investigated persons think it is unmeaning or not necessary.

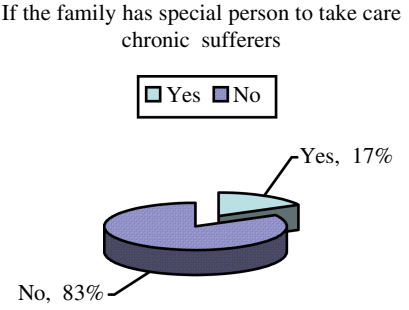


Fig. 5. Availability of Special caretaker in family

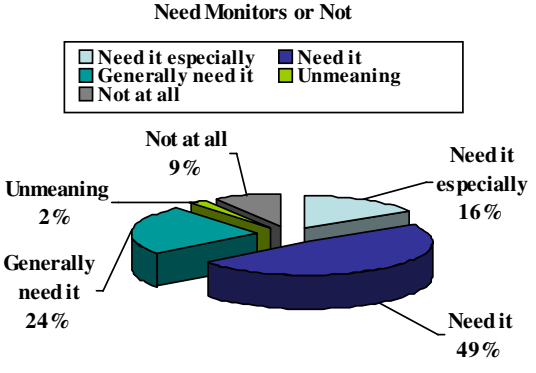


Fig. 6. People's subjective opinion on the necessary of the monitoring service

In allusion to the available parameter-monitoring and services by wireless healthcare network, the survey results are shown in Table.1 and Table.2. According to the results, we find that the requirement is different among different age levels. This is due to various life experience and self health status of different people. In general, the elder people hope their health condition can be more controlled than the younger people, as shown in Table 1. However, there is also commonality in different age levels. For example, because the high percentage of high blood pressure as shown in Fig. 4, both groups have a large requirement of blood pressure monitoring; and the

Table 1. Monitoring services required

Items	<30 years old	30~50 years old
Continual Monitoring in 24 hours	16%	15%
Customization/personalized	56%	56%
Results sent to hospital	17%	35%
The relatives obtain the results	20%	41%
Get instruction from doctors	20%	37%
Emergency treatment	22%	36%

Table 2. Physiological parameters to be monitored

Parameters	<30 years old	30~50 years old
Pulse	41%	33%
Blood pressure	72%	67%
Electrocardiogram	42%	48%
Heart rate	43%	41%
Blood sugar	53%	52%
Posture	15%	14%
Body temperature	25%	16%

blood sugar monitoring follows. And for the monitoring service, most people hope the service is customized and personalized.

4.2 Acceptability

As for the acceptability of WEHealth services in terms of health monitoring, as shown in Table.3, accuracy (60%), portability (52%) and the equipment's negative effect to body (42%), such as radiation, are the three most concerned items by most people. Therefore, these aspects should be paid special attention to during the industrialization of WEHealth systems and services.

Table 3. Worries about using wireless healthcare services and related equipments

Consideration	Attitude
Negative effect to body	42%
The leak of healthcare information	20%
Hard to use	15%
The price	36%
Will to take it	14%
Accuracy	60%
Timeliness	38%
Portability	52%
Transmission speed	9%

As far as time is concerned, 34% of the investigated people hope to use the WEHealth services and related products in the recent half a year; 40% hope to use the services in one year; 20% hope to use them within the next three years and only 6% want to use the services in the next five years or even later.

5 Discussion and Conclusion

An overall analysis of questionnaires reveals some interesting results and some problems we should be pay more attention to. The wireless applications used in healthcare industry in China has a huge market potential. People in modern big city such as Beijing, have a high acceptability degree of the innovative services and

related equipments of WEHealth. By the way, they also hope this innovation can partly solve current problem existing in Chinese medicine care system.

First, we summarize our investigation in technology aspect. The development of the system should comply with the requirement of people on WEHealth. For example, hypertension is the chronic disease suffering more Chinese families than diabetes and heart disease, etc. Therefore, we expect more acceptability of hypertension monitoring than other chronic disease monitoring when different WEHealth systems are promoted. Because of the high acceptability of people of WEHealth, government and related institutions should expedite the development of WEHealth services and the medicine care revolution. Besides, government should encourage the exploit and use of WEHealth in rural area, although this area is out of the investigation of this paper. This is because the percentage of mobile phone holders in rural area is getting higher and higher in recent years. However, during the extending of WEHealth, enterprise and manufacturer must pay more attention to accuracy, portability and safety mentioned above. The products should be "easy-to-use" and "friendly" enough to enable users to master the products as quickly and convenient as possible. For example, if some WEHealth service is implemented on mobile phones, the users generally could not endure a long access time due to the limitation of network data rate and capacity to the useful information.

Second, we analyze the problem in social aspect. In the whole structure of WEHealth, remote family monitoring is an important element. Investigation reveals that people, especially elderly people, will concern to where the monitoring information will be sent and who will deal with this information. In China, community clinics are not all-pervading and trustable as big hospitals. Hence, people can reassure only when the monitored message is sent to some relatively high level hospitals. To solve this problem, government should construct and develop community clinics or hospitals, at the same time, set up the trust of people to them. Another problem concerned by elder persons is that they hope the hospital deploying the remote monitoring services to be one of the hospitals with which their insurance covers.

In the spirit of developing and constructing reasonable WEHealth industry, the results of the survey can be used to make improvements to provide more convenient and useful services for people. Only based upon an exact and validated investigation, the technology developing and related strategy decision can be more conformable, correct and harmonic.

References

1. Kang, G., Liu, D., Ouyang, Y., Zhang, P.: Investigating the performance of IEEE 802.15.4 for medical applications. WSEAS Transactions on Communications 5(6), 1126–1134 (2006)
2. Poon, C.C.Y., Zhang, Y.T., Bao, S.D.: A norvel biometrics methods to secure wireless body area sensor networks for telemedicine and M-Health. IEEE Communications Magazine, pp. 73–81 (April 2006)
3. Ouyang, Y., Li, S., Chen, X., Kang, G.: Investigation and implementation on the advanced wireless medical registration solution in China. HCII 2007 (accepted for publication) (2007)

4. Pattichis, C.S., Kyriacou, E., Voskarides, S., et al.: Wireless telemedicine systems: an overview, IEEE Antennas and Propagation Magazine, pp. 143–153 (April 2002)
5. http://www.most.gov.cn
6. The application tutorial of the 11-th Five Year Plan of National Key Technology Supporting Project in China – the 14-th subproject: Area Cooperative Medical Service Demonstration Project
7. Maglaveras, N., Koutkias, V.,et al.: Home care delivery through the mobile telecommunications platform: the citizen health system (CHS) perspective. International Journal on Medical information 68, 99–111 (2002)
8. Varshney, U.: Patient monitoring u7sing infrastructure-oriented wireless LANs. International Journal on Electronic Healthcare 2(2), 149–163 (2006)
9. http://trochim.human.cornell.edu/KB/sampling.htm

Users and Trust: The New Threats, the New Possibilities

Kristiina Karvonen

Department of Computer Science, Helsinki University or Technology, P.O.Box 5400, 02015-TKK, Finland
Kristiina.Karvonen@tml.hut.fi

Abstract. In this paper, we describe the current threats and possibilities in the online environment from a user's perspective on basis of literary analysis and a set of user studies concentrating on online trust formation and expression. The work includes a review on the current level of understanding about online trust, what kind of research has been done and what the basic ingredients of trust might be in the future. By comparing the new results with results gained in 1999, we can see how the situation has been changing. The analysis is based on research conducted on trust between 1999 and 2007, utilising user studies, ethnography, cultural studies, aesthetics, and basic usability engineering as methods. We also include some discussion on the appropriateness of various methods available for studying online trust, and propose directions for new trust research as a field inside human-computer interaction.

Keywords: Security, human factors, trust, qualitative, user study, cross-cultural, identity, risk, perception.

1 Introduction

Understanding the real-world trust is crucial to understanding the actual security of any transactions online. It can be claimed that for Internet to survive, being able to grasp this difficult concept may be even more crucial than creating the actual technological solutions enabling these transactions [12],[26]. It has become a classical phrase of the usability of security to state that users are the weakest link in the security of on-line transactions [1]. This is probably true. Users will also remain so as long as they are expected to be able to make rational choices of whether an operation is secure and trustworthy or not on basis of misleading, hard-to-comprehend, or non-existing information. In fact, more often than not, there really are no bases for rational, argumentation-based decision-making from a user's point-of-view in an online situation requiring trust. In an un-trusted, open network like the Internet, a suspicious attitude might in fact present the right choice of attitude: It is to be expected that everyone will not be playing by the rules and be what they claim to be. This means also that the *identity* of oneself and others is hard to prove online.

Trust is also often based on false assumptions and false impressions in the real world, so why not also online. Users may decide to trust fortune-tellers and smooth-tongued salesmen, against their instinct and better interest [27]. Further, even the most decent of us are not always behaving like model citizens. Users have the need to put up "face" to others, and falsely emphasize their good qualities over bad ones. They

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 893–902, 2007.
© Springer-Verlag Berlin Heidelberg 2007

may also at times be too ashamed to admit they do not know something, and so on [15]. So, even in innocently intended interactions distortions and discrepancies in the integrity and authenticity of the information will appear – and it may happen implicitly and unconsciously [2].

A further issue is that even with good intentions applied with cautiousness and care, users still cannot really be trusted to flawlessly implement computer security policies. This is just the same as realizing that users cannot be trusted always to lock their car door or never to lose their keys [27]. It is only human to make mistakes sometimes; users cannot be trusted to do things precisely. People misunderstand and forget, but that is only human. A change in this inherently human quality of being erroneous is not likely to change any time soon, so it must be considered as a basic requirement for any security-related design. Preventing fatal errors from occurring in the first place is a top priority in security.

People's *needs* for security also seem to differ. For example, when it comes to privacy, e.g. [27] states that people have a complicated relationship with this phenomenon: When asked to pay for privacy, they often do not want to, but will rather do without – mainly because they may not understand how much is at risk. *Understanding the consequences of actions* is in fact quite hard for most online users [2], [16]. Also, what is *perceived as constituting a risk* may vary from one individual to the other. [29] and later e.g. [3],[8] have stated that people can be categorized into different groups according to the level of privacy they need. However, people are not always aware that there is something to lose when giving away their privacy. According to [20], privacy concerns have mostly focused on *prevention* of privacy breaches for the most, when in fact just as much attention should be directed towards raising the *awareness* of privacy and of *detection* of attempts to break one's privacy. Lack of awareness is, in fact, probably one of the biggest issues in breaking the user trust if privacy is lost as the result: Shneier goes as far as to claim that almost no-one realizes exactly how important privacy is in his or her life – before it is gone [27].

The work presented in this paper is based on user research, involving various methods provided by the framework of human-computer interaction, on online trust between 1999 and 2007. The initial trust studies were reported in publications [17][18][22]. The work at hand presents an update on the first study, reporting on findings on the recent developments in user attitudes towards online situations where trust is needed, and on the current online transactional behaviours among net users in Finland. The work is done in a project called "Trustinet". The main objective of this project is to study the trust provisioning in the Internet. As changing the current Internet infrastructure is difficult, the project will explore the use of overlay networks to build and manage trust relationships between service providers and customers over networks. In order to be fully successful, the human aspects of security must be considered from early on within the project. At the core of the work is concentration on the cognitive issues embedded in various trust-demanding situations.

2 Known Issues in Trust

Previous research on trust enables us to identify some major challenges in trust research. To start with, the payment process seems to have its own demands, different

from other online trust situations. Any disruptions in the payment process flow lead to quick decrease in the trust level, ultimately making the user withdraw from the transaction [10]. Depending on whether the user is acting in the role of user instead of consumer, where money becomes an issue, different things will matter, according to the role changes. E.g. [14] has identified clear differences in user behaviour between mere online shop browsing – "window-shopping" and making actual purchases – "actual shopping". Further, it is important to realize that both risk *aversion* and risk *gambling* are happening in the online world. Some users may be willing to gamble with their online relations, monetary, informational, or societal, giving away their private information in exchange to something that is tempting enough. According to [29] and [8], users differ in the privacy level they need and in their attitudes to controlling their own privacy. Most users are *pragmatists*, who care about the privacy and like to take measures in order to guarantee their privacy, but who would still be ready to trade-off the privacy for a good-enough incentive.

Discussion on the relation of trust and risk in the context of online environments was more or less initiated by [6], laying the groundwork for considering the risks of the Internet business. [6] identified the avoidable hazards which were likely to be found on the road towards realization of the Internet commerce. Risk perception has since then become one of the focal points of trust research. Researchers from Cheskin, one or the authoring companies of the influential Ecommerce Trust Study [10] have treated this issue in [9]. Their classification of different types of risks embedded in trust decisions is insightful. According to [9], when a person or corporation considers lowering the walls to act in the world, they have historically faced six types of risk, four of which often involve direct economic transactions. These are: 1. financial (risk of losing money or paying too much); 2. functional (risk of receiving the wrong or a malfunctioning product); 3. social (risk of embarrassment); and 4. physical (risk that we might be physically harmed). 5. In addition, engaging with the world puts us at risk in two other ways: 6. emotional (risk that one might be emotionally hurt by an interaction); and 7. identity (risk that others may impersonate us for financial or other types of gain). The goal of being able to trust, then, is risk reduction.

[26] set risk aversion in centerfold, and also state that users have practically no means either to perceive or evaluate the risks in each situation in a realistic fashion. The cure for better perception of risks that would, in effect, lead to enhanced trust towards the service, is offered using the traditional sources of trust outside the cyberworld for promoting trust. These include banks, insurance companies and other authorities, such as state offices. These traditional actors could find a new role as guarantors also in online transactions.

On basis of the user studies we run in 1999 [18], this would probably work: these actors, if online also, are already trusted more by the users than other online actors. Users in [18] were very trusting of their online banks, and would have hoped for all payments to be possible via a link to their online bank. Another interesting remark by [26] is that when the financial (or other, such as desirability) incentive grows, users are more willing to engage in transactions they otherwise would consider too risky. This feature came up also in [18] and [17].

3 The 2007 Update on Trusting Behaviour and Attitudes

This work presents the results gained from the interview part about the general attitudes of the users as expressed after first having been interacting with a security management tool.

3.1 The Study

In a user study on creating a usable security management tool for handling online identities reported in full in [19], 18 users were asked to discuss their current trusting behaviours in online environments in a post-test interview session.

The user study structure was thus very similar to the studies reported in [18], where users were first asked to use an online service and then discussed with on their online attitudes and behaviours.

The users represented age group between 19-39 years of age, mostly male, and one female. The users all had a rather technical background, being students or graduates from computer science from a technical university in Finland.

3.2 The Outcomes of the Interviews with Users

In comparison with the 1999 study, the users seemed more aware of their security. Most users reported awareness of security indicators on websites and claimed they were actively following them. These indicators included pictures of locks in the browser, changing colour of address bar and the added 's' in the URL, and certificate announcements, all associated with the SSL protocol usage that was trusted by all users in our study.

Most users reported that they were indulging in online transactions, which was a big change from the previous study, where only the minority of users were conducting ecommerce. However, users' trust towards the online service providers may not have grown: the users reported using several means to protect their assets online, such as using only sites they knew well, or only making payments via their bank's online services. Among further measures to fight against abuse were having multiple credit cards: one for offline and one for online purchases, with very limited credit limit on the latter in order to minimize risk.

Overall, even if users seemed to have learned to look for indicators of trust, and were involved in online transactions, the attitudes, the sources of information and amount of interest in security had changed surprisingly little. The users claimed they would not mind if someone would gain access to their personal e-mails, since "they didn't have anything to hide" – a claim they usually tend to abandon once the privacy gets breached. Further, security was, still, seen as a burden, and users were not really interested in it – they did not follow security news, and did not worry about bad things happening to them. Stories of others having been less fortunate were told, more or less in the same format as in the 1999 study.

4 The New Issues in Trust

How to enhance trust in the online environment? It seems the old issues remain, but new ones have emerged, too.

4.1 Trustworthy Design

Enhancing the visibility of security and trust-related concepts is an important issue. If we think of "firewalls" and "viruses" – these can be visualized, so people understand them since they have counterparts elsewhere. Small wonder products built around these concepts that are representing 'visible' or 'understandable' threats tend to sell rather well. There is a need to make also the rest of the security more visible/understandable, and to create concepts that epitomize the essential concepts of security to users in an easy to understand way.

In the future, it would be interesting to investigate further into what kind of simplicity would be right for creating trustworthy design. In the study by [21] it became obvious that what is simple may not be that simple (*sic*). Graphic designers often complain that usability experts always want design that is too simple, which they commonly interpret as boring. Simplicity in this sense is a kind of "stripped" simplicity – the design is stripped naked of all fancy features, colours, and flashy, moving objects. Is this what users really want? Or could there be a second kind of simplicity that they actually mean, "designed" simplicity – clear, and "clean" like the Swedish users wanted [17], but in a stylistic and beautiful way that does not lessen the pleasure provided? We think it is this latter form of simplicity that is asked for. However, it is a well-known fact that simplicity is not really desirable since it is not socially acceptable and because it does not reflect a self-image one could be proud of [23]. People want more: more buttons, more functionality, and they want to see themselves as "experts", above the average. They do not want to be labeled as those who use "simple designs". In any case, the importance of aesthetics in making trust decisions has become evident.

4.2 Access to Understandable Information

One of the greatest obstacles for securing the weakest link is lack of appropriate security information. E.g., according to a recent report by European Union on makings of Safer Internet [11], users would be quite interested in becoming more knowledgeable about security, but do not have access to appropriate information. Take any privacy statement or firewall attack report as an example – they are incomprehensible for the most users. As end result, most users will either ignore their safety or shy away from using the services provided, when they encounter such thus 'encrypted' announcements. The common example of gestalt psychology applies: when you don't know what you see, you can see anything: in darkness of the night, a lone traveler, vulnerable to attacks, can interpret the innocent shadow of a tree, twisting in the wind, as malicious attacker with bad intentions – and vice versa. The climate of the interaction dictates the connotation.

4.3 Emphasizing the Emotional Side of Trust

Trust is by its very nature built in part on cognition and on the other part on emotion. It has become commonplace in trust research to emphasize that there is another side to trust: the emotional, partly irrational side. To study such a topic, introducing personal psychology and consumer satisfaction is well grounded. This is, in fact, closely related to the recent major trend inside HCI: the study of feelings as an

integral part of this interaction. This trend has several sources. The previous concentration on the cognitive, rational factors in how humans behave was simply too narrow in scope. That paradigm was naturally born out of the earlier concentration on work environments, where computer were at first only used. Taking the affective side of human into account was a natural step, as the field as a whole has matured (c.f. [25] for ground-breaking work towards this expansion of the field in scope). Secondly, the study of HCI has left the office and entered the user's spare time, treating users of machines not only as workers but as individuals with personalities, likings and dislikings, and so on.

Studying the affective side of HCI follows naturally with these new directions. When the computers are used not only for work, but as tool for personal fulfillment, for socializing, for art-making – it is no longer enough to study just the cognitive side of the human participant of this interaction. The new applications may be immersive games or chats or dating services that by nature are designed to entertain and deal with feelings, more so than to appeal to the information processor inside the brain.

Even more, this information processing is no longer treated as a property of the brain alone, but intelligence needed in rational decision-making is constantly detected also elsewhere (e.g.[21]). This new approach is, however, built on the existing knowledge on 'traditional usability' of effectiveness, efficiency and satisfaction – this basic level of usability, consisting of things like memorability, learnability, and so on, is needed as basis for the emotional level to be built on in order for it to be possible and to work. Emotion-provoking, immersive experiences need the traditional aspects of usability as their groundwork in order to be able to fulfill their goals, i.e, to evoke feelings. The development from physical ergonomics some 60 years ago to today's analysis of the entertaining elements of computer games and online environments and virtual communities seems long yet natural and follows a logical path.

4.4 Understanding Perception of Risks

Defining different types of risk as in [9] seems quite useful from trust research perspective. Parallel to the analysis of the multi-faceted nature of trust, a similar analysis to risk – an undeniable factor in trust-formation – is recommendable. Using the approach by [14], the significance of the existence of some type of risk or their combination in a trust situation – a primary requirement for the need of trust in the first place – could undergo a similar analysis, where the sum effect of various types of risks could become calculable. First, we would need to analyse, what types of risks appear in what type of situations. For example, when conducting online transactions as in online shopping, risks number (1), (2), possibly also 3, and 6 from the presented list will become effective. On the other hand, being active in newsgroups or chat rooms seems to denote the existence of the risks (3), (4), (5), and (6): using these services may not be socially acceptable for some people, and depending on their topic – meeting an online 'friend' may also realise the physical risk, if one is unlucky. If the purpose of being online is to find a life partner, the emotional risk grows. Also, identity theft (7) is a serious risk in most situations.

Understanding the type of risk we are dealing with in each situation and linking them with the ingredients of trust might make this part of the trust decision-making measurable. Further, on basis of this understanding, a designer of an online service

might ultimately be able to help users notice the risks involved and act accordingly, thus leading to increased feeling of control of the service, and further to increased amount of trust towards that service. It remains to be seen, whether drawing such connections between all these ingredients is possible or worthwhile.

4.5 On the Social Nature of Trust

Placing 'the social' at the center of attention is a new phenomenon in the trust expression, but its roots are old. The upsurge of recommendation-based systems for formalising trust expressions and for making trust decisions is not surprising on basis of the trust studies conducted in 1999 [18] and in 2007 [19]. In fact, it seems the natural path for the trust-formation process. In the user studies, it has become evident that even in technical environments, humans trust, first and foremost, other humans – word-of-mouth, recommendations, or hearsay. In the most recent work of [7] we can also detect an interpretation of the current wide-spread interest on translating the real-world, often implicit, social networks into online explicitly expressed networks and modeling this. [20] state inhibiting users from recurring to established practices as one pitfall in failing to create working privacy solutions: stopping users from reusing the existing social ways to enclose or disclose private information about themselves to others in the online world also.

The social nature of trust has been extensively studied by [5][6], who states "that an examination of trust assumptions in security systems illustrates embedded values, and in particular these assumptions become embedded through technical simplifying assumptions that create social complexity, being among the first to investigate the social patterns of online contexts. The current approach aims for embedding existing social networks to the online environment and/or understanding the formation of new ones (e.g. [4]). This means treating users as *social cognizers* and breaking free from the initial topic of study of human-computer interaction – the individual user.

5 Conclusions

On basis of the research conducted so far by us and by others, we have been able to pinpoint few research lines that seem most fruitful. Where, then, should trust research go next? There are many areas that have been not researched extensively inside trust so far. For example, in order to fight against the incomprehensibility and ungraspability of the security concepts and terminology, it should be found out, how people talk about security. An inquiry, utilizing e.g. linguistic studies on which terms are actually used and how when people discuss about security, would benefit the building of more understandable user interfaces for security-prone applications that would be based on users' own vocabulary, thus promoting more trust in their users, feeling in control of something they can truly understand and relate to. Making also use of the mental models concept [23] is till today underrepresented in the usability of security studies. Not much, if any, research exists on this topic as yet.

A further area of study is aesthetic decision-making. There is some convincing evidence that in trust decisions, surface level aesthetic cues play an important, or even

crucial role for initial trust formation and also usability assessment that remains consistent and decisive also after extensive later use [28][21][24][17]. It seems this type of decision-making is valid in trust situations, but why? Is it possible to detect the enabling features on a perceptual or even neural level [21]?

Ubiquitous computing will present further challenges for trust research. Novel environments, e.g. netsurfing on the beach, the café etc. may affect the trust-formation principles and processes in unexpected ways. How do the trust requirements change when we enter the ubiquitous environment, or when RFID technology is used? Understanding and embedding the context in order to automatize and support the trust decision-making may be needed. In a way, in the future users may be "always online". What does it mean for trust? Can the devices used in trust situations intelligently analyse the contexts and go, say, to "safe-mode" when user is mobile in a condensed area, e.g. in the city center, or away from the device?

Another issue in makings of trust today is *identity*. E.g. Camp [5][6] has examined this questions of trust, risk and identity in identity management, voting systems and e-commerce, providing a quite usable identity taxonomy that could be tested with real users. The technology used, of course, presents its own definitions on what an identity consists of. Different network layers have their own demands for the ingredients of an identity. Further, the users have their ideas about identity, both online and offline, and they need to be taken into account when creating identity management solutions.

[26] interpreted the users' willingness to sometimes indulge in risky behaviour as denoting that users were perceiving the risk in such situation as smaller than it actually was. However, on basis of our user studies, it can be claimed that the users may in fact be well aware of the risk. Only, the risk is better endured, in hope of a possible gain, and users seem to be willing to give up on their normal 'trust level'.

This is equivalent with the "privacy gambling" concept, and could be phrased as a 'law' where there exists a correlation between the threshold value of tolerable risk and the financial incentive. According to such a law, the risk evaluations are not fixed but co-dependent with the possible gain: if the possible gain is perceived as greater than the possible negative effects of the risk realised and the damage it may cause, users are willing to take that risk. It would be interesting to be able to convert these risk evaluations computational somehow, in order to build a mathematical hypothesis on the co-dependence of risk evaluation and financial incentive. Measuring the level of trust felt towards an online vendor seems, however, to be hard.

Still today, there exists a huge gap between what users should know and what they in fact know about how computer security works. Also, no matter how we define what we mean by security, at the core of the issue lies another problem: there may be a great difference between actual and perceived security. What this means is that the users' ideas about what is secure and what is not might not have anything to do with the actual level of security of, say, a service on the Web. What users might perceive as risks and what they recognize as signs of trustworthiness, may be something quite different than any designer of the service might imagine. In part, this difference is due to the amount of ignorance: where there is no knowledge, there is room for superstition and false beliefs.

5.1 Future Work

We are currently in the process of finalizing the first round of user tests as in [18] in Finland. The next step will be to run the study in other Nordic countries in order to detect how the cultural differences detected in the earlier study still apply.

References

1. Adams, A., Sasse, M.A: Users Are Not the Enemy. Communications of the ACM 42(12), 41–46 (1999)
2. Ackerman, M.S., Mainwaring, S.D.: Privacy Issues and Human-Computer Interaction. In: Cranor, L., Garfinkel, S. (eds.) Security and Usability. Designing Secure Systems that People Can Use, O'Reilly, pp. 381–400 (2005)
3. Ackerman, M.S., Cranor, L., Reagle, J.: Privacy in E-Commerce: Examining User Scenarios and Privacy Preferences. In: Proceedings of the ACM Conference on Electronic Commerce, pp. 1–8 (1999)
4. Bonhard, P., Sasse, M.A.: Knowing me, Knowing you - Using Profiles and Social Networking to Improve Recommender Systems. BT Technology Journal, vol. 25(3) (July 2006)
5. Camp, L.J.: Digital Identity. IEEE Technology & Society 23(3), 34–41 (2004)
6. Camp, L.J.: Trust & Risk in Internet Commerce. MIT Press, Winter, Cambridge, MA (2000)
7. Castelfranchi, C., Falcone, R., Marzo, F.: Being Trusted in a Social Network: Trust as Relational Capital. Lecture Notes on Artificial Intelligence (LNAI). Springer, Heidelberg (2006)
8. Cranor, L.F, Reagle, J., Ackerman, M.S.: Beyond Concern: Understanding Net Users Attitudes about On-line Privacy, AT&T Labs-Research Technical Report TR 99.4.3 (1999) http://www.research.att.com/library/trs/TRs/99/99.4/
9. Diller, S., Lin, L., Tashjian, V.: The Evolving Role of Security, Privacy and Trust in a Digitized World. In: Jacko, J.A., Sears, A. (eds.) The Human-Computer Interaction Handbook - Fundamentals, Evolving Technologies, and Emerging Applications, Lawrence Erlbaum Associates Inc. Mahwah, NJ (2002)
10. Ecommerce Trust Study, Cheskin: ECommerce Trust Study. January 12, 1999. Cheskin Research and Studio Archetype/Sapient 31. (1999) http://www.cheskin.com/cms/files/i/articles//17__report-eComm%20Trust1999.pdf
11. Eurobarometer survey on Safer Internet http://europa.eu.int/information_society/activities/sip/docs/eurobarometer/eurobarometer_2005_25_ms.pdf
12. Friedman, B., Kahn Jr, P.H, Howe, D.C: Trust Online. Communications of the ACM 43(12), 34–40 (2000)
13. Fukuyama, F.: Trust: The Social Virtues and the Creation of Prosperity. Free Press, New York (1995)
14. Gefen, D.: Reflections on the Dimensions of Trust and Trustworthiness among Online Consumers. In: Gefen, D. (ed.) ACM SIGMIS Database. Summer 2002, vol. 33(3), pp. 38–53. ACM Press, New York, USA (2002)
15. Goffman, E.: The Presentation of Self in Everyday Life. New York, Anchor-Doubleday (1967)

16. Karat, C.M., Brodie, C., Karat, J.: Usability Design and Evaluation for Privacy and Security Solutions. In: Cranor, L., Garfinkel, S (eds.) Security and Usability. Designing Secure Systems that People Can Use, O'Reilly, pp. 47–74 (2005)
17. Karvonen, K.: The Beauty of Simplicity. In: Proceedings of the ACM Conference on Universal Usability (CUU 2000), November 16-17, 2000 Washington DC, USA (2000)
18. Karvonen, K.: Creating Trust. In: proceedings of the Fourth Nordic Workshop on Secure IT Systems (NordSec'99) November 1-2, 1999, Kista, Sweden (1999)
19. Karvonen, K., Komu, M., Gurtov, A.: Usable Security Management with HIP (to appear 2007)
20. Lederer, S., Hong, J., Dey, A., Landay, J.: Personal Privacy Through Understanding and Action: Five Pitfalls for Designers. In: Personal and Ubiquitous Computing, Springer, Heidelberg (2004)
21. Lindgaard, G., Fernandes, G., Dudek, C., Brown, J.: Attention web designers: You have 50 milliseconds to make a good first impression! Behaviour & Information Technology 25(2), 115–126 (2006)
22. Nikander, P., Karvonen, K.: Users and Trust in Cyberspace. In: Christianson, B., Crispo, B., Malcolm, J.A., Roe, M. (eds.) Security Protocols. LNCS, vol. 2133, pp. 24–35. Springer, Heidelberg (2001)
23. Norman, D.: The Design of Every-Day Things, MIT Press, London (1986/1998)
24. Norman, D.: Emotional Design: Why We Love (Or Hate) Everyday Things. Basic Books (2004)
25. Picard, R.W: Affective computing. MIT Press, Cambridge, MA (1997)
26. Salam, A.F, Rao, H.R., Pegels, C.C.: Consumer-Perceived Risk in E-Commerce Transactions. Communications of The. ACM 46(12), 325–331 (2003)
27. Shneier, B.: Secrets and Lies. Digital Security in a Networked World. John Wiley & Sons, USA (2000)
28. Tractinsky, N., Katz, A.S., Ikar, D.: What is beautiful is usable. Interacting with Computers 13, 127–145 (2000)
29. Westin, A.F.: Harris-Equifax Consumer Privacy Survey 1991. Atlanta, Equifax, Inc. (1991)

Ontology-Based User Preference Modeling for Enhancing Interoperability in Personalized Services

Ju-Yeon Kim[1], Jong-Woo Kim[1], and Chang-Soo Kim[2,*]

[1] Department of Infomation Security, Pukyong National University, Busan, Korea
{jykim,jwkim73}@pknu.ac.kr
[2] Division of Electronic, Computer and Telecommunication Engineering, Pukyong National University, Busan, Korea
cskim@pknu.ac.kr

Abstract. Personalized services that provide users with adaptive information filtered from heterogeneous and numerous information makes human-computer interaction efficient. Although personalized services have been researched in the various, heterogeneous information and users have raised the problem of defining a generic user model that would be used as a basis for recommending personalized information. Moreover, it is difficult to share the users' information. In this paper, we propose ontology-based user preference model as one of the solution to resolve the problems. In our model, user preference is described in User Preference Description Language (UPDL) with various domain ontologies. UPDL is a simple and OWL-based description language that allows users to describe their preference over domain ontologies. It provides not only rich expressiveness, reusability and flexibility but also the novel personalized services that can access and utilize not only user preferences in the own service but also in other services.

Keywords: Personalized Services, Interoperability, Ontology, User Preference Modeling, Domain Ontologies.

1 Introduction

Recently, a lot of researches on the services for supporting the ubiquitous computing are undergoing in the various areas with the growing interest on ubiquitous computing. Personalized service is one of the key technologies that provide users with adaptive information filtered from heterogeneous and numerous information [10]. It also makes human-computer interaction efficient, even in the restricted environments like mobile devices. Although personalized services have been researched in the various [1,3,10,12], heterogeneous information and users have raised the problem of defining a generic user model that would be used as a basis for recommending personalized information. Moreover, it is difficult to share the users' information.

In this paper we would like to propose Semantic Web approach as one of the solution to enhance interoperability among personalized services. Semantic Web is a

* Corresponding author.

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 903–912, 2007.
© Springer-Verlag Berlin Heidelberg 2007

technology to add well-defined meaning to information on the Web to enable computers as well as people to understand meaning of the documents easily. In this paper, we focus on user preference that is one of the important information of user profiles. For our approach, we propose ontology-based user preference modeling. We especially specify an ontology for describing user preferences. It can provide interoperability among heterogeneous personalized services by allowing user preferences to be described over domain (or service) ontologies.

In section2, we introduce related work on personalized services and Semantic Web. In section 3, our approach is briefly introduced. In section 4, we represent how to describe user preferences including User Preference Description Language (UPDL). In Section 5, we describe personalized service architecture based on our model. Discussion and conclusion is given in Section 6.

2 Related Work

Personalization dynamically adapts a system's service or content offer in order to enhance the quality of users interaction with the system. Having as its goal a closer customer relationship, personalization provides support to satisfy the needs, preference or goals of individuals and specific target groups [10]. For personalized services, systems depend on personal information describing user's characteristics, such as preferences, behaviors and other relevant facts. Therefore, personalized services should be represented well user's requirements, because user's needs dynamically change over time.

There are existing researches for personalization approaches in the various fields. MUSIC [3] extracts several features, and the information consists in databases by classifying its features in the smart office space. The architecture infers current mood of users and recommends preferred music and favorite genres to users through process of agents with gathered contextual information. Also, there are various applications in E-Commerce fields. Amazon.com is known for commerce communication with users in many e-commerce web sites. Tourism [12] provides a framework represented as XML-based user profile. The framework illustrates user's preference information by filtering LBS information through tourism scenarios. LBS [1] provides tourists with personalization, location awareness, interaction facilitation on handheld devices in order to provide various and convenient tourism services in the proposed CRUMPET system. The system enables agents to process personalized and context-aware support of mobile users with ontology-based tourist domain like hotels, restaurants, sights.

Recently, The Semantic Web technologies provide potentiality to enhance personalization functionality. The Semantic Web [2] is a technology to add information on Web to well-defined meaning, to enable computers as well as people to understand meaning of the documents easily, and to automate the works such as information search, interpretation, and integration. The automated agents and dedicate search engines can be a high level of automation and intelligence because the Semantic Web documents have a meaning that a computer can interpret, contrary to existing HTML-based Web documentations. The Semantic Web consists of Uniform Resource Identifier (URI), UNICODE, Resource Description Framework (RDF),

RDF Schema, and Ontology hierarchically. The RDF is an XML-based language to describe resources like images, audio files, or concepts available via the Web. The RDF model is called as a triple because it has three parts, subject, predicate, and object. The subject and object means a resource, and the predicate defines their relation. The ontology defines the common words and concepts used to describe and represent an area of knowledge. The Semantic Web documents must be represented as the RDF which defines common data of special domains based on an ontology which represents data relation. The contents on this structure can be searched, integrated, and inferred semantically. The systems using the Semantic Web have been developed in practical fields, and its effectiveness was also presented. Tanen et al. [13] have developed the Courseware Watchdog which is a comprehensive approach for supporting the learning need of individual in the environment of fast change using Ontology. Lee et al.[7] presented an architecture how a service flow can be automatically composed by syntactic, semantic, and pragmatic knowledge in medical services called ”Semantic Medical Services(SMS)”. Hunter et al. [5] presented a Web services-based system which they have developed to enable organizations to semi-automatically preserve their digital collections using ontology.

In personalization field, MyPotal [4] provides database-based personal semantic web information retrieval systems. Each provider describes their capabilities in WSCD(Web Site Capability Description) and is assigned a PSA(Provider Search Agent). Each consumer describes user’s requirements including his preferences. It is assigned a consumer search agent (CSA) and also has an intelligent unified user interface agent (UIA). This architecture shows that agents like CSA and PSA is searching and accessing tasks satisfied with user’s queries. FLAME2008 [9] provides Personalized Mobile Applications with user’s interests such as Olympic games, Shopping, Sightseeing, Entertainment. This system represents ontology-based user’s interests by considering user’s locations and contexts. Although these efforts have improved personalized services, there are several challenges. First, heterogeneous information and users have raised the problem of defining a generic user model that would be used as a basis for recommending personalized information. Next, it is difficult to share the users’ information.

3 Our Approach

Our goal of this work is to develop an ontology-based user preference model and profiling mechanism that supports the novel personalized services that can access and utilize not only user preferences in an own service but also user preferences in other personalized services. A distinguish feature of our profiling is to share user preferences with domain-specific (or service-specific) ontologies by providing interaction mechanism based on ontologies.

Fig. 1 depicts the conceptual service model of our user preference modeling for personalized services that is organized for the following two advantages. First, our modeling provides rich expressiveness of user preference by specifying them based on domain-specific ontologies. Because domain-specific ontologies are classified as the purpose and the feature of each domain, they can represent user preferences more correctly. Second, our mode provides strong interoperability among personalized

services. Because user preference is specified by OWL-based UPDL and domain-specific ontologies, if a term of a user preference in the user profile is one with terms in the ontology of personalized service used, the service can compute similarity from corresponding concept of higher level.

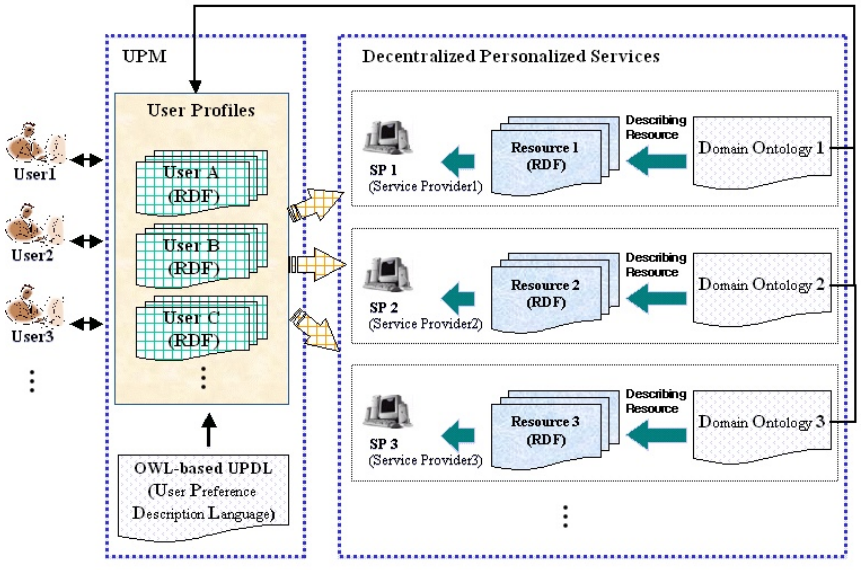


Fig. 1. Conceptual service model for ontology-based user preference modeling

The novel services are provided based on User Preference Description Language (UPDL) that is a simple and ontology-based description language that allows users to describe their preference over domain specific ontologies. The goal of the UPDL is to provide rich expressiveness for specifying the preference items and their weighting and to provide interoperability of such preference information. This goal can be satisfied with the use of the ontology that specifies the concepts of corresponding domains and their relationship, and supports the sharing of above information. In Section 4.2, detail of UPDL is described.

4 Describing User Preferences

This section presents ontologies for user preference and acquisition process for user preferences. User preferences are described based on UPDL ontology and heterogeneous domain ontologies each with their unique XML namespace.

4.1 Domain Ontologies

Domain ontologies define concepts for each domain and relationship between them. The domain ontologies can be specified in service providers or Standard

Organizations. Because taxonomies of contents are different as their domain, the domain ontologies help domain-specific information to be represented more exactly. In our work, these domain ontologies are referred by the resource of each service to classify them and the instances of UPDL ontology to specify the domain-specific user preferences.

Existing metadata and standards can be used to specify domain ontologies. SUMO[1] ontology is being created as part of the IEEE Standard Upper Ontology Working Group. The goal of this Working Group is to develop a standard upper ontology that will promote data interoperability, information search and retrieval, automated inferencing, and natural language processing. The MILO[2] ontology is a mid-level ontology that is intended to act as a bridge between the high-level abstractions of the SUMO and the low-level detail of the domain ontologies. In the field of computer Machinery, The ACM Computing Classification System (CCS)[3] can be used as a classification ontology. The full classification scheme involves three concepts: the four-level tree (containing three coded levels and a fourth uncoded level), General Terms, and implicit subject descriptors.

4.2 User Preference Description Language (UPDL)

User Preference Description Language (UPDL) is a simple and ontology-based description language that allows users to describe their preference over domain ontologies. The goal of UPDL is to provides rich expressiveness for specifying the preference items and their weighting and to enable the sharing of such preference information. This goal can be satisfied with the use of ontology which specifies the concepts of corresponding domains and their relationship and supports the sharing of above information. UPDL can describe not only the preference for a specific domain but the preference for other various domains. Therefore it can deal with the preference for concepts used in any domain dynamically, and enables the sharing of information by referencing ontologies.

User needs a way of flexible expression for user's preference to describe the interest concerning various domains and items. UPDL specifies the list for a user's various preferences and each preference separately. And because the relation between the preference list and each preference can be associated with the properties defined in UPDL, the flexible description of user's preferences is possible. Fig. 2 shows a simple example of user preferences for the paper retrieval service described based on UPDL and domain-specific ontologies.

In the example, Profile class is used to describe the list of various preferences for a user who associated with user's personal information and each preference information by two properties explained below. However the details of real preference is described by the instances in the Preference class.

- updl:personalInfo: This property is used to associate an instance of Profile class with the instances of Person ontology which express the personal information. In

[1] Suggested Upper Merged Ontology, http://ontology.teknowledge.com/
[2] Mid-Level Ontology, http://ontology.teknowledge.com/
[3] The ACM Computing Classification System, http://www.acm.org/class/1998/

Fig. 2, this property is used to associate Profile of 'UserA' with his/her instance of Person ontology (&person)' modeled to describe the information about the Person such as FOAF. The instance of Person ontology includes personal information such as id, name, and address etc.

- updl:hasPreference: This property is used to list various preference information and to associate corresponding instances of Preference class. In the example, this property is used to associate instances in the preference list with two preference classes, "UserA_InterestingField1" and "UserA_NotInterestingField1".

```
<updl:Profile rdf:ID="UserA">
    <updl:personalInfo rdf:resource="&person;UserA">
    <updl:hasPreference rdf:resource="#UserA_InterestingField1">
    <updl:hasPreference rdf:resource="#UserA_NotInterestingField1">
</updl:Profile>

<updl:Interesting rdf:ID="UserA_InterestingField">
    <updl:item rdf:resource="&papaer;SemanticWeb">
    <updl:weight>10</updl:weight>
<updl:Interesting>
```

Fig. 2. A simple example of user preferences for the paper retrieval service

Preference class describes the interesting information together with corresponding item and weight. Practically, user preference is described by two subclasses, Interesting and NotInteresting classes which describe the weights of items interested and not-interested respectively. Preference class describes the preference information using two properties below.

- updl:item: This property is used to associate items with a concept which is hierarchically classified in the domain ontologies. This property can be used to associate the concepts described in various ontologies together through URI (Unified Resource Identification). In the example, this property associates "SemanticWeb" concept of "paepr" ontology (&paper) that classifies research field hierarchically.
- updl:weight: This property describes the weight of preference. In the example, a value "10 is used as weights. These weights are computed through acquisition process for user preference. The detail of selecting weight is described in the Section 4.3.

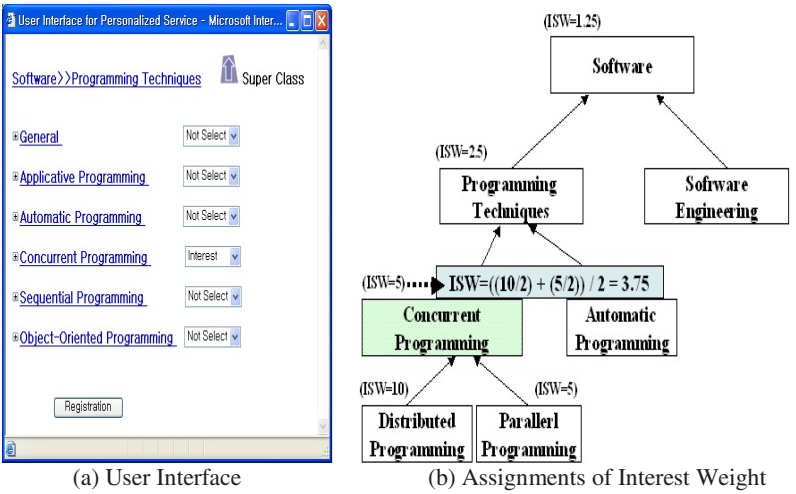
As we explained in this section, UPDL enables the description of personal preferences using ontologies in various domains and corresponding URI, and provides a way of flexible expansion of such preferences.

4.3 Acquisition of User Preferences

Based on the modeling, User Preference Manager (UPM) described in Section 5 provides the registration process with Web-based user interface. The registration process includes three sequence stages: 1) acquisition of user's interest information based on each domain specific ontology, 2) evaluation of weight for a item interested, and 3) updating of user preference profile. The acquisition process of user preference for user profiling is shown in Fig. 3 (a).

To collect user's interest, the concepts classified hierarchically by domain ontologies are provided through user interface. Once a user selects a preference concept, the UPM provides more specific (or narrower) information including concepts in the subclasses. And then the user can select the closest preference level to his/her preference information. Binary model is typically used for the user interest acquisition [11]. Behavior-based Binary class model provide for users to express the preference level. In Binary model, user's useful behavior pattern can be analyzed as a positive example and a negative example with machine-learning techniques.

Introducing two interest levels of Interesting and Not-interesting with the analysis of user preference information can solve the ambiguous selection problem in the system using multi-level of preference. It is an advantage that collection of user preference information in various domains is possible in the process of user preference acquisition without modifying the program for user interface.



(a) User Interface (b) Assignments of Interest Weight

Fig. 3. Acquisition of user preference

Acquired user preference information is described in UPDL with a corresponding weight. In our work, a simple assignment method is used. For items of the Interest class, the weight value "10" is assigned. Then the weight value of 50% less than it is assigned in its super class, and the process is repeated until root class. For items of the NotInterest class, the weight value "-10" is assigned. In the case of conflict among

interesting weights of subclasses, the average of half values of each subclass value is assigned. This weight assignment based on the hierarchical classification has very important advantage in the recommendation process.

5 Architecture of Personalized Services Using UPDL

We can enhance interoperability of users' preferences by building the User Preference Manager (UPM) that provides acquisition service, access service, and recommendation service as shown in Fig. 4. The Acquisition Module provides user interface that enable users to create and update his/her profile based on UPDL and domain ontologies. The Domain Ontology Registry is used to link the URI of each domain ontologies. The Access Module provides API (Application Programming Interface) that enable various personalized services to access users' profile. The Similarity Evaluation Module computes similarity from distance of corresponding concept of higher level in domain ontologies. The value is used in recommendation services in each personalized service as a factor for computing recommendation value.

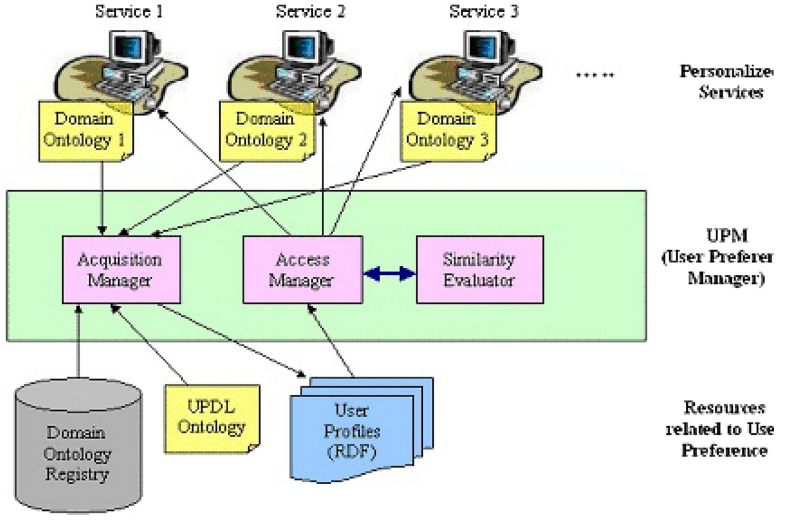


Fig. 4. Architecture of personalized services using UPDL

The modules in Fig. 4 were implemented using Jena, Joseki, and RDQL templates. Jena API is a Java application programming interface that creates and manipulates RDF documents, and Joseki is a Java client and a server that implemented the Jena network API over HTTP [6,8]. We can semantically search the instances of RDF documents through RDQL, a query language for RDF, which is Jena's query language [8].

6 Conclusion

Our goal in this work is to develop an ontology-based user preference model and profiling mechanism that supports the novel personalized services that can access and utilize not only user preferences in an own service but also user preferences in other personalized services. For the description of user preference, we specified UPDL that is an OWL-based preference description language. It allow user preferences to be described with domain ontologies. Based on the proposed user preference model, we realized the acquisition process as a module of User Preference Manager (UPM) and briefly introduce other modules of UPM.

Our approach can enhances human-computer interaction by recommending user adapted information from the user preferences specifed in other personalized services. An example scenario looked like the following: A user usually uses a paper search service to find his/her interesting paper about the field of 'ontology'. Now, if the user wants to use e-learning service, e-learning agent can recommend classes related to 'ontology' using user preferences specified in the paper search service. Even if e-learning agent does not find the same term with user preference in the ontology of e-learning domain, the agent can recommend by computing similarity from corresponding concept of higher level. Main contributions of our approach include:

- Interoperability: Our approach supports sharing personal preferences in various personalized services because personal preferences are described by the UPDL specified based on OWL.
- Rich expressiveness: The UPDL ontology provides more expressive and more flexible description mechanism for personal preference by allowing personal preference to be described over various domain-specific ontologies.
- Benefits of personalized services: Our approach provides advanced mechanism for personal preference profiling. Because our profiling approach is based on hierarchically classified ontologies, it is possible to recommend not only with explicit keyword but also with similar concept.
- Flexibility: Even if new domains are appended and domain specific ontologies are extended, user preferences are described in same method without modifying the acquisition process.

In our future work we will improve UPDL specification and User Preference Manager by studying more adaptive algorithms and methods. We will also experiment through realizing various personalized services.

Acknowledgments. This research was supported by the Program for the Training Graduate Students in Regional Innovation which was conducted by the Ministry of Commerce, Industry and Energy of the Korean Government.

References

1. Bara, S.B., et al.: Personalized and Location-based Mobile Tourism Services
2. Berners-Lee, T., Hendler, J., Lassila, O.: The SemanticWeb. Scientific American (2001)

3. Donghai, G., et al.: A Context-Aware Music Recommendation Agent in Smart Office. In: Wang, L., Jiao, L., Shi, G., Li, X., Liu, J. (eds.) FSKD 2006. LNCS (LNAI), vol. 4223, Springer, Heidelberg (2006)
4. Haibo, Y., et al.: An Architecture for Personal Semantic web information Retrieval System. In: WWW 2005, ACM, New York (2005)
5. Hunter, J., Choudhury, S.: A Semi-Automated Digital Preservation System based on Semantic Web Services. ACM, New York (2004)
6. Jeremy, J., et al.: Jena: Implementing the Semantic Web Recommendation. ACM, New York (2004)
7. Lee, Y., Patel, C., Chun, S.A., Geller, J.: Compositional Knowledge Management for Medical Services on Semantic Web. In: WWW2004., ACM, New York (2004)
8. McBride, B.: Jena: A Semantic Web Toolkit. IEEE INTERNET COMPUTING. IEEE (2002)
9. Nobert, W., et al.: Using Ontologies in Personalized Mobile Applications. In: GIS'04, ACM, New York (2004)
10. Riecken, D.: Personalized views of personalization, Communications of the ACM. vol. 43 (2000)
11. Sebastiani, F.: Machine learning in automated categorization, Computing Surveys, ACM (2002)
12. Shiju, Y., et al.: User Profiles in Location-based Services: Make Humans More Nomadic and Personalized. In: Proceedings of the IASTED International Conference on Databases and Applications (2004)
13. Tane, J., Schmitz, C., Stumme, G.: Semantic Resource Management for the Web: An E-Learning Application. In: WWW2004, ACM, New York (2004)

A Non-visual Approach to Improving Collaboration Between Blind and Sighted Internet Users

Ravi Kuber, Wai Yu, and Graham McAllister

Sonic Arts Research Centre & Virtual Engineering Centre
Queen's University Belfast, Belfast, BT7 1NN, UK
{r.kuber,w.yu,g.mcallister}@qub.ac.uk

Abstract. Assistive technologies have widened access to the Web, by providing a non-visual representation of page content for blind individuals. However, restrictions imposed by screen readers and Braille output devices have led to the development of mental representations of a web page, which differ considerably from the visual layout. As a result, both blind and sighted users find it difficult to perform joint web-based tasks. In this paper, we have applied a structured participatory-based approach to develop haptic sensations to assist the collaborative process. Preliminary results are reported and on-going development is discussed. It is anticipated that findings from the study will inform a haptic framework, which will assist developers when designing an accessible collaborative interface, using the benefits which the sense of touch can offer.

Keywords: Blind, collaboration, haptic, Web accessibility.

1 Introduction

There are two million people living with sight loss in the UK [8], many of whom use the Web for purposes of work, education and leisure. Screen readers and refreshable Braille displays are commonly used by individuals unable to rely on their functional vision, to access the contents of a page. However, these assistive technologies are known to impose restrictions on the browsing process, posing a challenge for users when attempting to develop an overview of content. While previous studies have mainly focused upon assisting blind users when independently exploring a web page, research has indicated that additional support is required to collaborate on web-based tasks with sighted peers [6]. Examples include working with colleagues to design web sites, and with telephone-based advisors to complete a series of web-based forms or to download and install software from a page.

Blind users often find the process of receiving directions from their sighted peers challenging, as it is difficult to locate or reach a particular target on an interface, using commands provided by an assistive device. As few screen readers and Braille output displays offer a visual marker, sighted users are often left confused as to a blind user's position on a web page. This can lead to dependence on the sighted user to complete

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 913–922, 2007.
© Springer-Verlag Berlin Heidelberg 2007

the task. To facilitate joint operation, visual and non-visual interfaces should ideally support the same model of the application interface, and their respective presentation should be synchronised [7]. However, the lack of positional information provided by assistive devices can lead blind users to envisage the layout of a web page as a vertical column-like structure, rather than the rich visually-dense representation perceived through sight [6,7]. Developing a sense of coherence between both visual and non-visual web interfaces would enable both groups to work together towards a common goal.

2 Related Work

Research has shown that non-visual technologies are beneficial to performing a collaborative task [1,9,10]. Crossan and Brewster [1] have discussed the design of a collaborative maze game for visually impaired and sighted children using haptic technologies. The introduction of a secondary haptic device is thought to aid awareness of position, promoting interaction with the virtual environment. Salnäss et al. [9] have asked visually impaired and sighted participants to locate, identify and manipulate objects using a haptic device. Results showed that visually impaired participants were able to form a mental representation to support orientation, and navigate through the environment guided using deictic references (e.g. “move left, right, up etc”) from sighted users.

It is thought that non-visual technologies can be used to assist blind and sighted users when performing a web-based task. Developing haptic sensations which can be accessed using traditional graphically-oriented browsing applications (e.g. Internet Explorer or Firefox) would enable blind users to gain the assistive feedback necessary to explore a web interface. At the same time, sighted users would be able to browse the Web in the manner which they are accustomed to.

3 Architecture

An assistive tool has been developed by our research group [11] which enables blind people to explore a web interface using the benefits offered by multimodal technologies. A force-feedback mouse is used for navigating around the interface, and its cursor position is constantly monitored by a content-aware plug-in. As the user moves within close proximity to an element located on a web page, appropriate haptic and auditory (speech and non-speech) prompts can be presented. Assistive guidance towards a particular target can also be provided.

Findings from a previous study have revealed that the multimodal interface provides a usable alternative to a screen reader for purposes of exploring the Web. Haptic technologies were found to be influential to enabling participants to locate interface objects (e.g. images and hyperlinks) and develop a spatial representation of page layout [11]. However, the evaluation focused on independent exploration of the

Web, rather than how participants would manage to work together on web-based tasks. Research has indicated that if an accessible user interface is developed, there is no guarantee that collaboration will necessarily be supported [10].

4 Approach

In order to design assistive haptic sensations which can support blind individuals when performing a joint task with sighted users, a structured method should be adopted. An approach proposed and detailed in an earlier publication [3] has been used to assist blind users when exploring the Web independently. The method has since been modified, and applied to the context of collaboration.

- **Step 1** – To observe blind and sighted users working together, a series of web-based tasks are designed and presented to pairs of blind and sighted participants. Blind users are observed whilst interacting with a multimodal interface and accompanying a haptic device to complete the tasks set. They are asked to discuss their thoughts regarding haptic feedback present on the interface. The session is video recorded and a hierarchical task analysis is also conducted. The aim of this step is to assess whether the haptic sensations provide enough support for the collaborative process, and to consider ways in which further assistance can be provided to support joint tasks between blind and sighted users.
- **Step 2** - On the basis of the findings from the analyses, a scenario relating to the haptic exploration of a web interface can be constructed. If findings from earlier observations have revealed that considerable time and mental effort is spent trying to perform a task with the haptic device, this finding should be included in the scenario for discussion. Prototypes should be developed by the researcher, to adhere to the needs of the users when performing a collaborative task.
- **Step 3 –** Assemble a design team (also referred to as a user group) consisting of haptic interface designers and blind screen reader users. The researcher should lead the group through the main issues covered by the scenario. The aim of this step is for the group to verify whether the prototypes serve their purpose, and provide further feedback to strengthen them. Alternatively, the group can suggest other design ideas. To do this, participants should think about the role that haptic sensations would play on the interface, and then consider special qualities that touch has to offer. If participants make a design suggestion to aid collaborative work, they should consider whether the sensation would work in conjunction with mappings defined for use on the web-based interface. These suggestions can be developed and presented to successive design teams.
- **Step 4** - Assemble further user groups comprising of blind screen reader users and haptic designers, and present them with the same scenario and prototypes resulting from steps 2 and 3. This stage aims to iterate designs until they reach an acceptable standard. At this point, representations can be placed within the assistive framework.

4.1 Step 1: Observations and Analysis of Blind Users Performing Tasks with a Haptic Device

Haptic sensations were developed using an earlier application of the methodology [4], to represent images, hyperlinks, image-hyperlinks and borders (Table 1)[1]. These mappings were then integrated into the assistive tool described by Yu et al. [11].

Table 1. Haptic mappings used on web interface

Elements	Haptic Mappings
Images	A slightly raised enclosure effect is used to encase the visual border of the image. A weak spatial texture is applied to the image's interior.
Image-hyperlinks	A spring effect is used to direct the user towards the relative centre of an image-hyperlink. A spatial texture effect (the same texture applied to images) is applied to provide awareness of the element.
Hyperlinks	A spring effect is used to direct the user towards the relative centre of a hyperlink. A weak periodic wave effect to alert presence of the hyperlink body.
Page border	A lowered enclosure effect is used to provide a border around the HTML content. This allows the user to explore the page, and perceive its boundaries. To leave the page, the user is able to push-out on to the raised border. A vibration is encased around the border, to provide additional positional awareness for the user.
Page background	No haptic sensation should be used to represent the background, as the other mappings on the page should stand out. The location of text on a web page should not be communicated through haptic technologies.

Fourteen blind volunteers, aged between 18 and 65, were recruited for the study. All participants used a screen reader to access computers, as they experienced limited levels of functional vision or complete occlusion.

The aim of the task was to investigate whether both blind and sighted users could work together towards a common goal. The scenario of working with a telephone-based advisor to perform an installation of a web-based plug-in was presented. Each partnership would be required to move through a series of six unfamiliar web pages, to locate and install the plug-in. Each blind participant was asked to use the force-feedback mouse to explore the assistive interface, and locate certain targets (e.g. hyperlinks, image-hyperlinks) present within each page. Directions to each target were communicated verbally by the sighted user (played by the evaluator). Each selected target would open a new page, where another target would be located. This process would continue until the final target on the sixth page had been located, signalling that the plug-in had successfully been installed. The sighted user was unable to view movements made by the blind users, relying on their verbal feedback to communicate their current location on the page. A time limit of ten minutes was imposed to explore each page. An extended version of the procedure is presented in Kuber et al. [4].

[1] Since the experiment, a further session was conducted to iterate findings, discussed by Kuber et al. [3].

Findings from the study revealed that blind participants were able to perceive and identify objects presented on the web interface. They were able to locate a particular landmark on a web page (e.g. large image or hyperlink), and use this as a basis of communicating their spatial position to the sighted user. In turn, the sighted user would provide directions to another fixed reference point within the page. This process would continue until the blind user had located the actual target (average 61.09 seconds). All participants managed to complete the series of steps needed to install the web-based plug-in. By exploring the page, participants could use the spatial feedback presented to construct a mental map of object layout. Through the course of the experiment, blind participants' levels of confidence appeared to grow using the haptic device to follow spatial directions. It seemed that blind and sighted participants had developed the shared mental understanding of interface layout, discussed by Salnäss et al. [9].

Many participants adopted interesting strategies to support exploration. Borders were often used as a 'ruler', providing a means to move in a straight direction to another section of the page. Some participants would use the border as a fixed reference point, to try and ascertain their relative position on the interface, by gauging the distance to the nearest corner. This technique did not always provide an effective representation as participants noted that making small movements with the mouse would often correspond to larger movements on the interface.

Participants seemed to be able to locate and target larger hyperlink targets quite effectively. The attractive feedback was found to be helpful, as it reduced the opportunity to deviate from the central region of the hyperlink. The weak vibration present on the body of the link helped to provide awareness of position on the interface. Participants were generally able to use this sensation to align the mouse over the hyperlink to select it.

Areas for Improvement. Analysis revealed that participants experienced the following issues relating to spatial orientation and navigation on a web interface:

- **Assessing current position on the page** – Participants had no quick method of assessing their current spatial position on the interface (e.g. whether they were at the left or the right-hand side of the page), and communicating this information to a sighted user. Some participants were forced to spend time attempting to isolate an object (landmark) to try and convey current location.
- **Moving along a straight path** – Participants stated that they could navigate in the direction specified by the evaluator when using the force-feedback mouse. However, they were unsure as to whether they were adopting a straight path or deviating in another direction altogether.
- **Locating elements on the interface** – In some instances, certain elements proved difficult to locate on the interface. Problems mainly occurred when attempting to find small targets (e.g. textual hyperlinks and image-hyperlinks) within large empty regions of a page, due to the lack of surrounding feedback. Time and mental effort would be required to explore the interface to isolate these targets.

It is thought that haptic sensations could be developed to support the three areas identified.

4.2 Steps 2 & 3: Developing and Presenting Scenarios and Prototypes

A scenario was constructed highlighting the needs of Alison, a blind user with no functional vision, performing a task in conjunction with her sighted colleague, Laura. As Laura sits at a desk facing Alison, she is unable to view actions made by Alison when exploring the Web. The task selected was to access a time management system to enter in weekly hours. It was chosen because one blind participant from an earlier study had explicitly mentioned in an interview, her problems using a screen reader to enter data into a web-based form. As a result, she would ask sighted users for assistance. However, due to the limited level of coherence between visual and non-visual presentation, she would often ask sighted participants to complete the task on her behalf. Prototypes to assist navigation and orientation were developed by embedding haptic sensations within static HTML pages. These were based on design ideas provided by blind screen reader users and haptic designers in an earlier study [3]. An excerpt from the scenario is shown below:

> *Unsure about her location on the interface, Alison makes a keystroke, turning on an optional haptic effect allowing her to* ***identify her current position to communicate to Laura****. Using the haptic sensation provided, she is able to ascertain that she is located at the left-side of the page. This allows Laura to redirect Alison towards the button located towards the bottom of the page, to open the web-based form.*

The scenario was presented to the design team comprising of one blind haptic interface designer, one sighted haptic interface designer and one blind auditory interface designer. The design team would be able to listen to the scenario, interact with prototypes, and suggest whether sensations designed could target the three issues identified. They could then discuss amongst themselves how to improve the haptic feedback, or alternatively suggest new ideas for design. It was hoped that any findings would not be solely constrained to filling out web-based forms for time-management systems, but also apply to other collaborative contexts of use.

Assessing Current Position on a Page. Findings from Kuber et al. [3] had revealed that to assess the current position on a web page, participants had suggested providing two distinctive sensations, one at the either side of the page. Both effects should merge into one another towards the centre of the interface. It was thought that if the user was unable to locate a distinctive landmark to communicate his/her relative position on the interface, he/she could use the optional sensation to provide awareness of location. For purposes of testing, participants were asked to explore an interface with a spatial texture at the left side of the page, and a temporal vibration at the right, as these effects were thought to be the most distinctive from one another.

Participants were presented with a web page containing a number of images and the haptic orientation effect designed (Fig. 1). They were asked to identify the position of each image on the interface, and count the number of images present. Results showed that both blind users were able to accurately identify the location of each of the images positioned on the page. However, the effect was thought by participants to be inappropriate for purposes of detection, as it was a time consuming method to identify object location. This was due to the amount of haptic feedback present on the page. One participant stated that the application of the effect on the

interface would result in confusion, and not add any value to the interface. It was thought that more benefit could be gained from using the border as a means for referencing location. Discussion between members of the group inspired the blind haptic designer to suggest an idea for improving the effect. It was thought that four different haptic effects could be placed along the page border, one to mark the position of each quadrant. If the user wished to identify location on the page, he/she could move to the border and gain positional feedback in a less intrusive manner. This sensation would be mocked-up and tested with a new user group. The use of notches embedded within the border was briefly discussed by the group, as a method for locating position within the page. However, the idea was later rejected, as they thought that counting notches along the border would be too time-consuming, defeating the purpose of the effect.

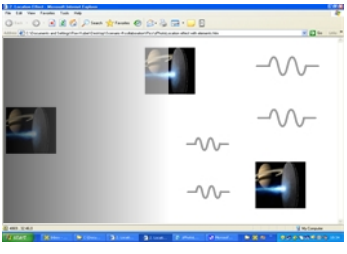


Fig. 1. A strong spatial texture present on the left-hand side of the page, reduces in intensity as the user moves rightwards. Similarly, a strong temporal vibration at the right-hand side of the page weakens as the user moves leftwards.

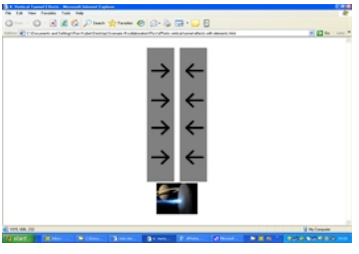


Fig. 2. Vertical haptic tunnel created with 20 pixel channel width. As user moves downwards through the channel, he/she is able to locate the element positioned that the bottom of the page.

Moving along a Straight Path. Participants had previously discussed the use of tunnels or grooves, as a method of moving in a line without deviating from the path [3]. Haptic tunnels were developed using the findings from Langdon et al. [5], where spring effects are embedded in each wall, separated by a channel width of 20 pixels. As the cursor passes over a tunnel wall, a spring force pulls the cursor to the inside wall [5], providing awareness of the path to be followed (Fig. 2).

As participants began to navigate through the tunnel, they tended to move from side-to-side hitting the inner walls, until they became more familiar with the path. All members of the design team thought the effect was appropriate for use, providing a fairly non-intrusive method of moving along in a straight direction. The blind haptic designer stated that she could feel enough resistance against the walls, enabling her to identify the limits of the channel. When asked how the effect could be improved, suggestions were made by both haptic designers to reduce the size of the springs, and improve the offset values to ensure that the user could take a clearer path through the channel. The blind audio designer thought that this improved effect would also be useful for guiding beginners (like the character from the scenario) from a search box to a submit button, or to move to areas of interest, aiding independent exploration, in addition to collaborative work. Participants felt that the tunnel would be better than

other attractive effects discussed, as this effect allowed them to be in control of movement, rather than being propelled to a particular target located in another region of the page.

Locating Elements on the Interface. Findings from Kuber et al. [3] had revealed that periodic waves increasing in intensity when approaching an object, would provide additional guidance to a user. Gentle effects attracting the user to an element located in close proximity had also been discussed. In the current study, participants were introduced to two sets of effects; periodic waves increasing in intensity and constant motion effects designed to move the mouse to the intended location. The effects surrounded small images, rather like a haptic aura. Both effects were designed to provide a subtle form of directional assistance towards the image.

Participants generally found both sensations quite useful to provide guidance to the intended targets. Both blind participants in the group found the directed motion more effective than the periodic wave effect. This was because directed motion would physically guide the user to target the area of interest. However, the sighted haptic designer felt that there was more information present within the periodic wave increasing in strength, which could enable the user to determine the relative proximity towards the element. The other members of the group suggested that periodic sensations would conflict with other periodic wave mappings (Kuber et al., [3]). To gain consensus on the most appropriate effect, both would need to be tested with successive user groups, to ascertain the respective benefits and weaknesses for purposes of locating a small element on an interface.

4.3 Step 4: Evaluating and Refining Sensations

From the basis of findings, prototypes will be improved and implemented in the multimodal plug-in. Further sessions will be conducted with design teams, to assess whether the sensations suggested provide benefit to blind users when collaborating with sighted peers. In order to ensure that sensations are refined to a sufficient standard, designs can be evaluated against the same scenario, and against additional scenarios exploring other contexts of use (e.g. jointly booking a ticket or installing a web-based plug-in). More comprehensive testing over a longer period of time will reveal how effectively feedback can be used to aid the collaboration process. Sensations would then be placed in the assistive framework accordingly.

5 General Discussion

The design approach has been used to isolate some of the problems faced by blind users, when attempting to work on web-based tasks with sighted users. Haptic technologies have been shown to be vital to aiding blind users to access interface objects and visualise layout of a page. Blind users were able to locate landmarks which could be used to convey relative position to sighted users. Both parties were also able to communicate in terms of a common language, which could not always be achieved through use of a screen reader.

Winberg [10] has stated that the non-visual presentation of the interface should be sufficient to let the blind user play an active role in the collaborative process and help to solve problems. Even though findings have showed that blind and sighted users were able to develop a shared mental representation of the interface, further observation and analysis of blind users working with a haptic device revealed that additional navigational and orientation support would be required for purposes of collaboration. As a mouse offers a predominantly visual-centric approach to exploration, blind users prefer to make use of keyboard shortcut commands in conjunction with screen readers and refreshable Braille displays. However, when a blind user interacts with a device such as a mouse, additional assistance is needed to make the precise movements required, and monitor the relative cursor location on a web interface as unintentional and unnoticed rotations can be made [2]. The results of our experiments confirm the benefits of locational effects (auras) and haptic tunnels to assist the navigation process. Attractive forces have been shown to be useful to guiding users towards a small target. However, the user should be able to resist forces if desired. In terms of the haptic tunnel used, Langdon et al. [5] thought that the freedom allowed by the presence of the wide channel could possibly be a hindrance to motion impaired users experiencing difficulty navigating in a straight line. However, the width of 20 pixels was thought to be adequate for blind users new to a mouse, as it provided them with the directional information necessary to follow the channel and complete the task. Participants did not feel constrained by this effect. The testing of the haptic effect to assess current position on the interface, has revealed that the overuse of haptic feedback on the interface can be counter-productive, and may reduce the quality of the subjective browsing experience. Sensations should be usable and should support the user in his/her intended task.

The importance of context has been emphasised by Winberg [10] when studying collaboration between blind and sighted users. Issues such as location, motivation and the experiences of people should be taken into account when the task is being designed. In the study, a real world situation of contacting a helpdesk advisor was recreated, and findings were observed. The scenario provided a useful means to evaluate prototypes and solicit design ideas for assistive feedback. The user groups were generally able to provide suggestions even if they had limited experience of a screen reader or haptic technologies. The participatory-style approach enabled both haptic interface designers and blind users to discuss and reflect upon each other's ideas, bringing the strengths of their own knowledge and experiences to the design table. In the participatory-design process, it is important to remember that end users are not expert designers. Users may make design suggestions, but these may not necessarily be appropriate or feasible for use on an interface. Testing is essential to ensure that suggestions fit their purpose, and work in conjunction with other mappings.

It has been acknowledged that there are restrictions on the information that haptic modality can provide. The haptic modality is limited in bandwidth, and devices are one-point in nature, limiting the extraction of attributes. Haptic feedback can be used to provide the majority of spatial and navigational information on an interface. However, it is not always appropriate to represent certain features (e.g. to quickly assess current position on a page), which may be better conveyed through another modality such as speech or non-speech audio.

6 Conclusions and Further Work

This paper has applied a structured approach to develop assistive haptic feedback to the context of a collaborative web-based task. Results have shown that haptic sensations are able to provide the spatial and navigational support needed. Application of the method has shown that design ideas can be developed and tested, targeted to the needs of blind Web users. Discussion can be generated amongst haptic interface developers and blind screen reader users to suggest new ideas which can be evaluated accordingly. As a next step in the research project, it appears pertinent to refine and test the haptic mappings designed. Further experiments, observing blind and sighted users performing tasks in a more natural environment would serve as a better indicator of assessing the merits and weaknesses of haptic feedback on a page.

References

1. Crossan, A., Brewster, S.: MICOLE: – Inclusive Interaction for Data Creation, Visualisation and Collaboration. In: Hands on Haptics Workshop CHI'05 (2005)
2. Jansson, G., Pedersen, P.: Obtaining Geographical Information from a Virtual Map with a Haptic Mouse. In: Proceedings of International Cartographic Conference (2005)
3. Kuber, R., Yu, W., McAllister, G.: Towards Developing Assistive Haptic Feedback for Visually Impaired Internet Users. In: Proceedings of CHI'07, ACM Press, New York (To appear 2007)
4. Kuber, R., Murphy, E., McAllister, G., Strain, P., Yu, W.: Evaluation of an Assistive Web Interface: A Collaborative Approach. Internal Technical Report (2007)
5. Langdon, P., Hwang, F., Keates, S., Clarkson, P.J., Robinson, P.: Investigating Haptic Assistive Interfaces for Motion Impaired Users: Force Channels and Competitive Attractive Basins. In: Proceedings of Eurohaptics'02, pp. 122–127 (2002)
6. Murphy, E., Kuber, R., McAllister, G., Strain, P., Yu, W.: An Empirical Investigation into the Difficulties Experienced by Visually Impaired Internet Users. Universal Access in the Information Society (To appear 2007)
7. Mynatt, E., Weber, G.: Nonvisual Presentation of Graphical User Interfaces: Contrasting Two Approaches. In: Proceedings of CHI'94, pp. 166–172. ACM Press, New York (1994)
8. RNIB Good Web Design. Available: http://www.rnib.org.uk/xpedio/groups/public/documents/code/public_rnib003460.hcsp
9. Sallnäs, E.L., Bjerstedt-Blom, K., Winberg, F., Severinson Eklundh, K.: Navigation and Control in Haptic Applications Shared by Blind and Sighted Users. In: McGookin, D., Brewster, S. (eds.) HAID 2006. LNCS, vol. 4129, pp. 68–80. Springer, Heidelberg (2006)
10. Winberg, F.: Supporting Cross-Modal Collaboration: Adding a Social Dimension to Accessibility. In: McGookin, D., Brewster, S. (eds.) HAID 2006. LNCS, vol. 4129, pp. 102–110. Springer, Heidelberg (2006)
11. Yu, W., Kuber, R., Murphy, E., Strain, P., McAllister, G.: Novel Multimodal Interface for Improving Visually Impaired People's Web Accessibility. Virtual Reality 9(2), 133–148 (2006)

Idea Movement of Aging Citizens: Lessons-Learnt from Innovation Workshops

Jaana Leikas

VTT Technical Research Centre of Finland,
P.O. Box 1300, FI-33101 Tampere, Finland
jaana.leikas@vtt.fi

Abstract. The demographic revolution, rapid aging in Europe has raised the creation of innovation market as one of the most important elements in the European strategy. During 2006 VTT Technical Research Centre of Finland carried out a campaign called the Idea Movement of Aging Citizens, where mobile service ideas were collected from aging citizens for open use. The ground for the campaign was the fact that the aging persons are the best experts of their own needs and should thus have more say in the development of new mobile services. Around 4300 ideas were collected in 13 innovation workshops in Finland. About 750 aging citizens participated in these workshops. This paper presents experiences and lessons-learnt of using innovation workshops as a method for activating aging people to work as idea generators for mobile services.

Keywords: User-driven innovation, aging citizens, mobile technology.

1 Introduction

The rapid development of information society in Finland as well as in the whole of Europe has led into a situation where mobile technology is becoming a part of everyone's life – including the aging people [12, 16]. People become increasingly dependent on their mobile devices. The services for a nomadic user are based on the assumption that every person has a mobile phone, which of course statistically is about to be true. Despite of this, mobile solutions are most often developed for the needs of younger people, who are "on the go" and have an increasing array of technologies at their disposal for keeping in contact with each other. For older users, the mobile phones are too small and complex, in the worst case decreasing older people's ability to communicate with their friends and families, not to mention the ability to benefit from mobile services provided by the society.

It has been argued that Finland has succeeded in combining the welfare society and competitiveness [2]. This is probably true. However, in Finland we are living in a society with increasing needs of the citizens. Ageing of the population is proceeding with high speed, among the highest in Europe. Our aging people are not a homogenous but a very diverse group. This diversity of people over 60 presents a challenge to researchers as aging people come from a wide variety of backgrounds,

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 923–931, 2007.
© Springer-Verlag Berlin Heidelberg 2007

histories and experiences and their demands and desires for technology are likely to reflect diverse personalities, educations and outlooks [4].

A core phenomenon in our modern society is individualisation. The lifestyle and set of values of people are becoming differentiated. One can no longer point out an archetypical pensioner. As people are expressing more and more individual needs the ability of the society to respond to these needs with universal Design-for-All-services is weakening. Universal services, i.e. services accessible to everyone, are designed to masses, standardised and regulated [13, 14]. Now the need for one's own individual voice has become a basic need for people's wellbeing, and customising services according to the client's needs is a major challenge. [11.]

The market success of a new mobile device is influenced by the consumer's total user experience of the product or service, starting with its image prior to purchase [7]. Thus, the design of new technology for the aged should be extended to cover more than what is *needed* by the older population and to explore what these people may *want* [10]. In order to do this we have to find more person-centred approaches and include user emotion and response to technology into this work. Involving users in service innovation brings new possibilities and also gives life to the service development process. It demands the use of new methodologies for creating dialogues with aging and older people to facilitate discussion about technologies [E.g. 4, 1 and 9]. Innovative strategies for enabling this dialogue include the use of innovation workshops - presented in this paper - where aging persons with no technical background invent new ways to benefit from mobile technology.

2 Idea Movement of Aging Citizens

The central characteristic of the information society is the ability to make innovations, i.e. to produce novel and original products and services in order to compete in global markets. Closed innovation, a traditional way of creating new products and services, is based on the idea that in order to profit from R&D we have to do everything ourselves: to discover new ideas, develop them and to ship them. This strategy does not, however, meet the needs of companies of the new millennium who seek for competitive advantage. They have to find new ways for creating innovations. Opposite to closed innovation is *open innovation*, where companies are challenged to efficiently utilize intellectual capital regardless of its source. The main principle of open innovation highlights that external R&D can create significant value and we don't have to originate the research to profit from it. However, internal R&D is needed to claim some portion of that value. [3, 5, 12.]

During autumn 2006 VTT Technical Research Centre of Finland carried out an outside innovation campaign called the Idea Movement of Aging Citizens, where mobile service ideas were collected from aging citizens for open use [6]. The campaign was carried out together with The Finnish Pensioner's Federation, Centre for Innovations for Independent Living (ISAK) of North Karelia University of Applied Sciences, and The Creative Tampere program of the City of Tampere, who all supported the campaign by offering facilities for the workshops.

Ideas for mobile services were collected in 13 innovation workshops. About 750 aging citizens participated in these workshops and produced around 4300 ideas. These

original ideas were then published for open and free use to anyone in business, research, government, education or other in February 2007. The ground for the campaign was based on user-driven design; on the fact that the aging persons are the best experts of their own needs and should thus have their say in the development of new mobile services.

3 Innovation Workshops

13 innovation workshops were organized around Finland with about 750 aging participants (Table 1). The age of the attendees varied from 50 to 89 years. About 87% of the participants were 61-80 years old (Table 2). In 5 out of 13 workshops the participants were specifically invited to produce ideas for new mobile services, whereas in 8 workshops they had come around the country to spend their holiday in a holiday resort (Perniö) and attended the workshop out of curiosity or in a hope of a free cup of coffee. Some participants attended the workshop hoping that they would get training on how to use a mobile phone. Almost each participant had a mobile phone but few had used a computer or the Internet.

Table 1. Number of innovation workshops, participants and collected ideas

Innovation Workshop	Venue	Number of Participants	Number of Collected Ideas
1	Tampere	53	490
2	Tampere	84	680
3	Perniö	39	250
4	Perniö	55	353
5	Perniö	76	358
6	Nurmes	5	30
7	Lieksa	40	165
8	Joensuu	34	202
9	Perniö	96	480
10	Perniö	72	400
11	Perniö	44	252
12	Oulu	84	280
13	Perniö	65	341
		Total 747	**4281**

Table 2. Breakdown by the age of the participants

50-60 years	**61-70 years**	**71-80 years**	**81-90 years**	**No ref.**
4.4%	45.6%	41.2%	7.3%	1.5%

A systematic brainstorming method was used to carry out the workshops. Each workshop took 1,5 – 2 hours, with settling down, introduction, innovation work and discussions (Fig.1). After a motivating introduction the participants started to create

ideas individually and document them on a special form. This took 10 to 20 minutes. The participants were introduced with the basic idea of the innovation work, where familiar ideas first tend to pop up from our memory, and after the feeling that "I can't figure out anything anymore" our brains are finally capable on creating new, innovative ideas. This introduction was found especially important, as most of the participants were not familiar with the brainstorming method and needed encouragement to innovate.

After the individual work, a short discussion was carried out in order to share ideas with others and to create an open and criticism-free atmosphere. After the discussion and a coffee break, group working sessions were carried out, where 3 to 4 persons in each group first shared their original ideas with each other and then created new ideas together.

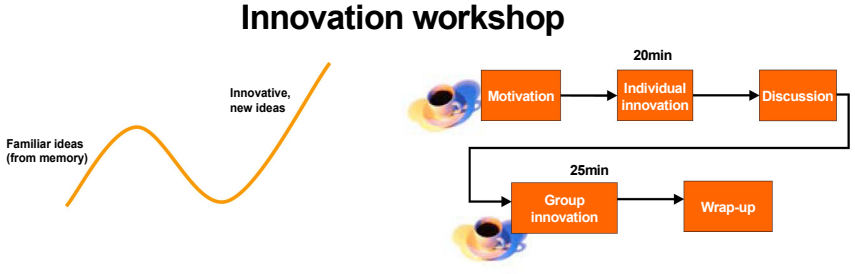


Fig. 1. Structure of innovation workshops

The individual work resembled exams at school, as people concentrated on writing down their ideas. During the group work sessions the atmosphere was clearly more relaxed, and a lot of humour was used, also. The moderator had a central role in the workshops. She introduced the idea of the workshop to the participants and was responsible on carrying out the brainstorming sessions within the time limit. Visualisations of the purpose of the event were found important: they helped people comprehend the concept of "a mobile service", the goal of the campaign and their own role in it. In some events there were musicians and choruses of elderly persons to create a cheerful feel.

Usually there were 1-2 assistants needed in the workshops to deliver and collect the documents. The assistants also participated in the group sessions chatting with different groups and encouraging people to express their ideas.

4 Aging Persons as Idea Generators

The amount of collected ideas was about 4300. As there were altogether 750 participants in the campaign, statistically the participants created an average of 5.7 ideas per person. However, the ability to produce ideas varied a lot between persons. The participants represented wide variety of backgrounds, histories and experiences, as well as attitudes towards new technology. Jarvis has studied older people's

responses to new experiences, and introduces three common ones: harmony seekers, the sages and the doers [8]. In the innovation workshops it was possible to discover similar types of participants, though the line between different types is somewhat vague.

The *doers* live an active life. They are physically fit and enjoy their retirement time. In the workshops it was possible to find out two types of doers: There were "the inquisitives", who were curious and interested in learning new things and technology and produced a lot of ideas. Some of them were also "jokers" who wanted to joke about everything, helped creating a cheerful atmosphere and produced the funniest and wildest ideas.

The *sages* are selective, choosing carefully the experiences that they will ignore. Having no sufficient interest to new technology, the sages in the workshops were somewhat critical towards innovation. In this group there were "the quitters", who expressed that they were not able to learn anything new anymore, and were happy to leave all innovation to younger people. Some elderly ladies who believed that technology was only men's business, belonged to this group, too. In this group there were also "the cheapskates" who valued every idea through its price and felt that most ideas would probably be too expensive to use, anyway.

The *harmony seekers* want to be at peace with the world and reject the potential learning experience if they cannot presume on the world around them. In the workshops "the technophobics" belonged to this group. These people expressed a concern that they would not be able to use any mobile service unless it was very simple, and were concerned that if they innovated something new the mobile phones would become even more complex to use. In the harmony seekers there was also a small group of "hideaways" who did not want to participate in the discussions. Usually these people hadn't had much education and were not accustomed to express themselves in writing or take part in group work. For these persons, attending an innovation workshop was probably too demanding, as the method required ability to make notes and actively participate in group discussions. As one lady in a workshop wrote: "My head is empty – could the mobile phone help with it?"

The time reserved for the workshops proved to be quite short for innovation for the participants. Aged people need much more time to accustom themselves to what they have been told than young people, especially if the subject is new for them. Introducing the aim and the structure of the innovation workshops took much more time than in a group of young people. Young people are used to grab the pen and start writing as soon as they have been told what to do. Older people, however, need time to digest the information and, in many cases, the information must be repeated several times (Fig.2).

In our workshops it was typical that when the time reserved for the workshop was nearly used, many participants were just about to warm up. When they had the chance to tell their ideas to others they opened up for new ones. This was also paradoxical: the participants were starting to warm up towards the end of the workshops but at the same time they found it tiring to sit down for such a long time. Having some easy gymnastic exercises together between the sessions helped concentrate on "brain gymnastics", and the participants enthusiastically carried out these exercises. Coffee breaks between the sessions brought also relaxation and enjoyment and paved way for criticism-free ideas.

Fig. 2. Individual work session in the innovation workshop

20 minutes was too long time for many persons to work individually. After a few workshops we decided to change the time slot of individual innovation into 10 minutes and increased the time of common discussions. This way the participants were able to at the same time share and nourish new ideas without getting too tired.

The participants created ideas for all walks of life, from activities of daily living to entertainment (Fig.3). The ideas can be categorised as "excellent", "interesting", "conventional", and "not a mobile service" ideas. Conventional ideas were produced most of all, and ideas for personal wellbeing and safety were the most popular topics.

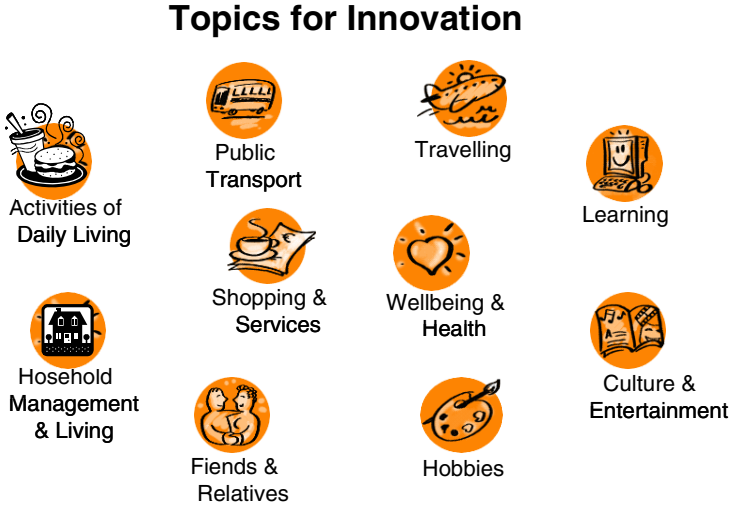


Fig. 3. Main topics for innovation

The participants created also wild ideas, e.g., related to finding a suitable partner. These ideas came mostly from the persons who attended a workshop during their holiday in the holiday resort.

5 Discussion

The Idea Movement of Aging Citizens was an open social innovation that offered aging people a possibility to attend in the society and feel social equity. Aging persons expressed ideas that were distinct experienced needs, closely related to everyday life, and not technology-driven. It would be possible to realize a considerable amount of the ideas with existing technology. The aged really enjoyed participating in the brainstorming sessions where they could share not only their new ideas but also their hopes and needs with others. The workshops also gave them a new learning experience and a possibility to give their share as consumers and individuals in strengthening the sustainable way of living.

The participants signed out a paper where they gave VTT the permission to use and refine the ideas they had created. Typically, the aging citizens were not too concerned about IPRs. It was found important, however, to publish the names of all the persons who participated in the innovation workshops on the Innovation Movement's web pages, and this way tell that these people had created the initial ideas. Also, if a specific idea is found especially brilliant in future, it will be illustrated and published in an idea book with the name of the "mother" or the "father" of the idea.

User-driven innovation is cost-effective hard work. Although the workshops were carried out in a relaxed atmosphere, one can hardly say that creating ideas was purely wild and creative activity. On the contrary, most of the time focused idea generation was tough work, with humour as a necessary flavouring in it. It was found out that the success of the workshops for the aging persons depends highly on the skills of the moderator and how (s)he understands the purpose of the research. It is her/his task to encourage and motivate people to create all kinds of ideas, without any criticism. (S)he has to be, at the same time, determined to carry out the interactive session and create an unreserved and sociable atmosphere. The moderator has to understand group dynamics and be sensitive for the group atmosphere. (S)he has to have ability to empathise the group members and thus have a neutral but encouraging attitude towards the statements of the participants. The moderator also has to be able to guide the dialog during the group discussions, and have discreet control so that (s)he can encourage all participants to take part in the discussion.

Difficulties in using current services still seem to be the main obstacle for the active aged to implementing mobile services in their everyday lives. Many good ideas as well as the usage of present mobile services are doomed to fail due to low usability of the services, from the design of mobile phones to incomprehensible operation logics. In this respect "Design for All" or "Universal Design" has not yet reached mobile devices and the needs of the aging population. In each workshop the participants expressed a lot of complaints on the user interface and navigation, as well as the size of the display, the buttons and the actual phone itself. They told that instead of fancy multifunction phones - that they find too difficult to use - they would rather like to use simple and big enough phones without any cameras, emails or music

possibilities. Combining the challenges of handheld devices with small displays and buttons and the need for the ease of daily use should thus be the foundation for any mobile service for the aged.

The Idea Movement of Aging Citizens gained a lot of publicity in Finland. Along with newspapers and national TV news, different organisations of senior citizens got interested in the campaign. Impartiality of the main player (VTT) was a benefit in organizing the campaign and gave credibility to it.

6 Future Work

VTT will continue the Idea Movement of Aging Citizens in the spirit of open innovation. The collected ideas will be refined and developed to service concepts in different co-design councils and focus groups. Actors in this future innovation network will be technology researchers and developers, service providers and aging people together.

Our workshops were organised mainly for persons in their seventies. For this second round of innovation it is necessary to invite the active aged in our society, people of the age of 50-65, who are relatively wealthy, relatively fit, and professionally skilled, and spending well-deserved free time. This group of work-dedicated people now want to express themselves and seek for more individual way of living. These peoples' needs are becoming differentiated as the material basic needs are relatively well fulfilled.

A detailed analysis and processing of individual ideas will be the main task in the refinement process. The main questions for this work will be: How to efficiently analyze a large set of ideas and how to generate real businesses out of mobile service ideas? And finally, a major challenge will be again to harness the creativity of aging persons to co-create new mobile-centric service concepts with ambient intelligence that would address their (future) everyday needs.

Acknowledgments. The author would like to acknowledge The Finnish Pensioner's Federation, ISAK and different civic organizations that participated in the arrangements of the innovation workshops.

References

1. Battarbee, K.: Co-Experience. Understanding User Experiences in Social Interaction (Dissertation). University of Art and design. Publication Series of the University of Art and Design Helsinki A 51. Helsinki (2004)
2. Castells, M., Himanen, P.: The Information Society and the Welfare State. The Finnish Model. Oxford University Press, Oxford, UK (2002)
3. Chesbrough, H.W.: Open Innovation. The New Imperative for Creating and Profiting from Technology. Harward Business School press, Boston, MA (2003)
4. Dickinson, A., Dewsbury, G.: Designing Computer Technologies with Older People. Gerontechnology 5(1), 1–3 (2006)
5. Hamel, G.: Leading the Revolution. Harvard Business School Press, Boston, MA (2000)
6. Idea Movement of Aging Citizens: htpp://www.idealiike.fi

7. Isomursu, P.: Personal Content Lifecycle in Mobile Product Design. In: Repo, P., Koskinen, I., Grönman, H (eds.) Naturalizing of Innovations. [Innovaatioiden kotiutuminen]. Annual Report of the National Consumer Research Centre 2006. Edita Prima Oy. Helsinki 2006. (In Finnish) (2006)
8. Jarvis, P.: Paradoxes of Learning. On Becoming an Individual in Society. ossey-Bass Inc. Publishers, San Francisco, California (1992)
9. Magnusson, PR.: Benefits of Involving Users in Service Innovation. European Journal of Innovation Management 6(4), 228–238 (2003)
10. Midden, C.J.H.: Motivation is a Process. Gerontechnology 5(4), 253–254 (2006)
11. Mokka, R., Neuvonen, A.: The Individual's Voice [Yksilön ääni. Hyvinvointivaltio yhteisöjen ajalla]. Report 69 of the Finnish Innovation Fund - Sitra. Edita Prima Oy. Helsinki 2006. (In Finnish) (2006)
12. Pekonen, O., Pulkkinen, L.: Social Capital and the Development of Information and Communication Technology [Sosiaalinen pääoma ja tieto- ja viestintätekniikan kehitys]. The Committee for the Future of the Parliament of Finland. Technology assessment report 11, the Secretariat of Parliament 5/2002. (In Finnish) (2002)
13. Von Hippel, E.: Democratizing Innovation. The MIT Press, Massachusetts Institute of Technology, Cambridge (2005) http://web.mit.edu/evhippel/www/democ1.htm
14. Weiss, S.: Handheld Usability. John Wiley & Sons, Ltd, England (2002)
15. Whitney, S., Keith, S.: Active Aging through Universal Design. Gerontechnology 5(3), 125–128 (2006)
16. WHO (World Health Organization), Active Aging: A Policy Framework (2002) Available online http://whqlibdoc.who.int/hq/2002/WHO_NMH_NPH_02.8.pdf

Accessing e-Government Services: Design Requirements for the Older User

Lorna Lines, Oluchi Ikechi, and Kate S. Hone

Information Systems & Computing
Brunel University
Uxbridge, UB8 3PH, UK.
{lorna.lines,kate.hone}@brunel.ac.uk

Abstract. Access to e-Government services, such as welfare, housing and financial assistance is typically achieved through accessing, completing and submitting paper-based forms. Current UK e-Government strategies indicate that such service access forms will be made available online. With the aging population considered as a primary user of these services this paper promotes the need to establish older adults' online form design requirements. Building on previous research [see 8], the current investigation validates and extends previous online form design requirements for an aging user group. The paper opens with an overview of current UK e-Government initiatives and promotes the investigation of older adults' online form design requirements, not only to assist this vulnerable user group with independent living but also to support social inclusion.

Keywords: Older adults, online forms, e-Government.

1 Introduction

A rapidly increasing aging population has led to concerns relating to UK government service provision [3]. As the UK's aging population increases [see 10], it is likely that so too will the demands on local government authorities in the provision of welfare, housing and financial services. At present, access to such services is typically by means of paper-based form completion and submission. However, e-Government initiatives in the UK indicate that electronic access to service provision should be made available to all citizens [see 6]. Although, in the past, internet technologies were considered as a young person's pursuit, rapid uptake of internet technologies by older adults has been predicted [see 2]. Thus online form completion and submission is likely to become a more frequently utilised method by older adults to obtain welfare, housing and financial services.

This paper supports the introduction of online forms for service access and their use by older adults. However, although highlighted briefly by form designers in the past [see 1], older adults' interaction design needs are often overlooked [4; 5]. Thus increases in the aging population, the uptake of internet technologies by older adults and the current UK e-government initiatives, presents the investigation of online form design as a timely topic for inquiry.

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 932–940, 2007.
© Springer-Verlag Berlin Heidelberg 2007

Building on the preliminary findings of Lines et al. (2004), whereby user requirements for the design of an online financial assistance form for older adults were presented, this paper sets out to establish the reliability and generalisability of these requirements to a wider set of e-Government service access forms.

The paper begins by introducing the benefits of online form completion and submission for older adults and discusses the design implications in relation to cognitive, sensory and physical declines that are frequently observed as a product of the natural aging process. Following this, a brief overview of the Lines et al. (2004) study will be provided and the online form design requirements obtained by this investigation are presented. Extending this previous work, the current investigation and findings will be discussed in light of previous work. The paper concludes with a revised set of design requirements and planned future research directions are offered.

2 Benefits of Online Forms for Older Adults

Online forms, at a conceptual level, are likely to yield a number of benefits for the older user. This section considers the benefits of online forms in contrast to the difficulties presented to older adults by paper-based forms. These benefits are considered in relation to three main activities; form access, form completion and form submission.

2.1 Form Access

Personally collecting paper-based government service provision forms from post offices or local government offices may be difficult, if not impossible, for older adults who present mobility impairments. For those who cannot obtain paper forms in this way, a telephone request for a paper form may be made, yet a delay in receipt of the form will be incurred due to postal delivery services.

Accessing e-Government service provision forms via the internet allows the older adult to access the form 'anytime, anywhere', as long as access to the internet is available. With a predicted uptake in internet technologies by older adults [2] a greater number of older adults are likely to have an internet connection available in their own home, or internet services accessible via libraries, day centres or in the homes of family members. Thus the older adult is no longer constrained by physical collection or telephony services that may hinder their access to service provision forms

2.2 Form Completion

Older adults may experience difficulties when completing paper-based government service provision forms. Based on anecdotal commentaries offered throughout our investigations, older adults typically request assistance during form completion. Assistance is sought for a variety of reasons, such as declines in manual dexterity that creates poor handwriting skills, visual declines that are exacerbated by the frequent use of 'small print', and question and answer clarifications so that appropriate data is provided. Often carers or family members provide such assistance during the form completion stage. However, this in itself may cause the older adult unnecessary

anxiety as requests for government service provision often require the disclosure of sensitive, personal and financial information for service eligibility assessments. Older adults may not be comfortable sharing this information with carers or family members.

By completing an online service provision form online, these difficulties and concerns can be reduced and/or eliminated. Appropriate adoption of assistive hardware devices for data entry can support manual dexterity declines; text can be manipulated to support visual declines, and; help functions could provide further information where questions are unclear.

2.3 Form Submission

Although guidelines exist [see 9] at present there appears to be little consistency in the online form submission process presented by UK local government agencies. Some local government agencies present the form online, yet there is no function to support the online submission of the form. In these cases, forms can be completed online, yet must be posted/delivered to the local government agency offices. Other forms require the older adult to print the form, complete it by hand and then return by post/hand delivery. Else, online submissions are supported by the local government agencies' technological infrastructure.

Paper-based form submission can create unnecessary difficulties for the older adult. Once again, mobility declines may be exacerbated by posting the form back to the government offices. Thus older adults may rely on carers or family members to post/deliver the form. However, this may cause unnecessary delays.

When information is omitted or incorrectly presented in a paper based form, the form is returned to the older adult for resubmission. This action then creates a delay in the service provision process. Allowing older adults to submit the form online may assist in detecting data omissions and/or inaccuracies simply through the use of a validation script that can run automatically at the point of submission; highlighting the errors/omissions and requesting 'corrected' information before submission is permitted by the system. This, although unlikely to eliminate the return of forms entirely, is likely to reduce the number of forms returned by the local government agencies, thus presenting a potential administrative benefit.

Using online mechanisms for delivery, in cases where forms require adjustments to be made by the older adult, the form can be returned electronically, once again reducing the amount of time required to complete/resubmit the form.

3 Online Forms for Older Adults: Design Challenges

Although some difficulties relating to accessing, completing and submitting paper-based forms can be easily identified when age related declines relating to the cognitive, physical and sensory ability are considered, there is little specific reference in the literature demonstrating how these difficulties may be overcome. Therefore the present study, building on the work recorded in [8], seeks to validate the earlier findings and extend the scope of the previous research by considering not only forms relating to financial support, but forms that serve as the access point to housing and

welfare government services. Given the lack of research activity in the area of online form design, it is not clear whether the additional forms will be supported adequately by the older adults' design requirements documented in [8] and reported below.

4 Previous Research Summary

This section provides a brief overview of the earlier investigation reported in [8], The methodological protocols adhered to during the investigation are followed by an examination of the key design features recommended by older adults (and subsequently evaluated via prototype implementation) for the design of an online, e-Government 'Housing and Council Tax Benefit' form.

To determine the strengths and weaknesses of the design of a paper based 'Housing and Council Tax Benefit' form prior to online implementation, eight older adults were recruited to participate in semi-structured interviews. Interviewees were selected based on their age (over 60), some knowledge and experience of the internet and the lack of severe cognitive, physical and sensory declines; other than those declines that would be expected as a product of the natural aging process.

Semi-structured interviews were deemed an appropriate methodology for use with an aging user group based on the observations reported by [7]. The interview opened with discussion concentrating on the interviewee's computer/internet use and their experiences of accessing, completing, and submitting both paper-based and online e-Government forms. During the interview, interviewees were requested to complete a local government 'Housing and Council Tax Benefit' form. They were also encouraged to employ the 'think-aloud' technique; whereby strengths and weaknesses of the form and the difficulties experienced during form completion are articulated by the interview while conducting this task.

By doing so, the strengths and weaknesses of the paper-based e-Government form design and current online form design were elicited. Having employed a thematic analysis technique on the data captured, design requirements were put forward and subsequently implemented into an online version of the 'Housing and Council Tax Benefit' form. These design requirements were positively evaluated by the interviewees during a follow-up, user evaluation session employing the same methodology. The following section considers these design requirements for the online 'Housing and Council Tax Benefit' form in further detail and provide the basis for the present investigation.

4.1 Older Adults' Design Requirements

From the user requirement capture and subsequent evaluation of these design features implemented within a prototype online 'Housing and Council Tax Benefit', five key design requirements (general and online specific) were identified and are summarised below:

1. *Form Layout:* Extra spacing between questions and between answer boxes is recommended to overcome difficulties related to poor vision and question and answer comprehension.

2. *Simplification of Question Structure:* Simplification of question structure is also recommended to avoid creating excessive demands on the older adults' cognitive load.
3. *Question Completion Assistance:* Pop-up messages and hyperlinks appearing next to the appropriate question are suggested to serve as a help function to further clarify or simplify a question and its completion.
4. *Additional Information:* In addition, information sources required to complete the form, for example, pension details, should be presented at the beginning of the form to avoid delay.
5. *Data entry:* By using keyboard and mouse devices handwriting legibility difficulties due to declines in manual dexterity are overcome. This is likely to reduce the need for the issuing organisation to return the documents for correction. In addition, by employing a data entry validation system, any obvious data entry errors of omissions can be identified prior to form submission, again reducing the need for the form to be returned and corrected.
6. *Form personalisation:* To avoid answering unnecessary questions, the online form should be 'personalised' based on the data entered by the user. For example, when marital status is entered and the user is a widow, no questions relating to the user's deceased spouse should be presented.
7. *Online form submission:* Submitting the form online will avoid the difficulties associated with posting/delivering the form and/or arranging for a carer to post/deliver the form. Online form submission will be of great benefit to those presenting mobility deficits. In addition, online submission of the form may lead to a quicker processing time achieved by the issuing organisation.

These design requirements were used to direct the current investigation reported below.

5 Validating and Extending the Online Form Design Requirements

The method for this study remained the same as [8] reported above. Once again, eight older adults, aged over 65, presenting cognitive, sensory and physical declines associated with the natural aging process and eligible for local government services were recruited. All participants reported some experience of the Internet. Whereas in the previous investigation a single form for accessing housing and council tax benefits was used, this study employed a financial service [Housing Benefit & Council Tax Benefit Form], a welfare service [Warm Front Grant] and a housing service [Housing Register New Starters Application Form] access form. The latter replicated the form used in the earlier investigation to build continuity and for direct comparisons to be made between findings.

As in the previous investigation, semi-structured interviews employing the think-aloud technique were conducted. Each interviewee was asked to complete each of the paper-based forms and the difficulties encountered, alongside positive features of the paper-based forms, were discussed to elicit the older adults design requirements for online versions of these government service access forms. The order in which the

paper-based forms were completed by the interviewees was counterbalanced, using the latin square technique to reduce order effects.

5.1 Key Findings and Discussion

Table 1 (below) presents the results from this investigation. The majority of the design requirements elicited by [8] are supported by this study (indicated by 'Yes' responses) and are directly applicable to the online design of the extended range of service access forms used here. Differences occur between the earlier study and the present study in relation to form layout and data entry requirements (indicated by 'No' responses).

Table 1. Validation of the design requirements elicited by Lines et al. (2004)

Design Requirements	*Financial Form*	*Housing Form*	*Welfare Form*
A) Form layout	No	No	No
B) Simplify question structure	Yes	Yes	Yes
C) Question completion assistance	Yes	Yes	Yes
D) Additional information	Yes	Yes	Yes
E) Data entry	No	No	No
F) Form personalisation	Yes	Yes	Yes
G) Online form submission	Yes	Yes	Yes

The earlier design requirement relating to form layout suggested that extra spacing between lines of text and answer boxes would be beneficial for the older adult. Although not supported by the present investigation, we feel that it would be unwise not to promote the need for 'appropriate' spacing.

Clearly, the current study revealed that an appropriate amount of spacing was offered by the paper-based forms for the interviewees recruited for this investigation. This finding is particularly surprising as the original recommendation was based on the 'Financial' service access form that was employed during both investigations and thus it is suggested here that difference between the two groups of interviewees', such as visual acuity and familiarity with form completion may have created such an effect . These between-group differences clearly require further investigation.

In addition, the earlier 'data entry' design requirement highlighted the need for automatic checking and validation of the data entered by the older adult. Although deemed unnecessary at this stage by older users, it is argued here that such a feature would be beneficial as a 'safe-guard' for eliminating/reducing the need for

government administrators to return forms and delay the process of service access provision. This feature is further supported by current design practice in financial, insurance and purchasing industries within their online environments.

Although the older adults were informed at the beginning of this form completion activity that this investigation was not a test of their intelligence, admitting to forgetting to respond to a question or provide incorrect information may have been perceived as a socially uncomfortable admission and a direct reflection of their personal abilities.

Interestingly, additional design requirements, to those elicited during [8] can be offered by the current investigation and documented in Table 2, below. These extended online form design requirements for older adults include; Bullet pointed instructions, logical information groupings, justification for personal/sensitive information, greater information relating to data security, a help and assistance function and a 'save and return' feature. Overall there appears to be a large amount of consistency between the design requirements offered by the older adults' across the service access forms considered here.

The first design requirement relates to the use of bullet points to indicate instructional content. This was justified by interviewees' in terms of a mechanism to support the identification of an instruction (e.g. how to complete the form) and a question.

Secondly, the interviewees' also suggested that information should be grouped in a logical manner and that they should not need to refer to previous answers that they had provided in the early stages of the form completion process in order to answer a much later question. This created confusion for some of those participating; perhaps creating increased demands on cognitive load as their attention shifted between questions and previous answers.

Table 2. Extended design requirements

Design Requirements	*Financial Form*	*Housing Form*	*Welfare Form*
Bullet point Instructions	Yes	Yes	Yes
Logical information groupings	Yes	Yes	Yes
Justification for personal /sensitive questions	Yes	Yes	Yes
Security information	Yes	Yes	Yes
Help and assistance feature	Yes	Yes	Yes
Save and return	Yes	Yes	No

Thirdly, the older adults' felt as though some justification could be provided in relation to questions that required the disclosure of personal or sensitive questions, such as their financial or health status, as they resented providing this information without knowing the value of this disclosure. The non-disclosure of this information could lead to data omissions and the return of the form to the older adult by the local government administrator if this was considered an essential piece of information.

Seemingly related to this, was the fourth design requirement of the provision of greater promotion of the security of the information that they provide; concerns raised related to who would see the paper-based forms and, if implemented on line, how could they be sure that their data would remain confidential.

Similar to the requirement offered by [8], the fifth design requirement found here, relates to that of a help/assistance feature. During the form completion activity many of the interviewee's required help that was not readily available with the form. When this was discussed in detail, there were suggestions that in an online form, a 'formal' help page could be offered in addition to the previously recommended 'pop-up' message or hyperlink next to a question. The interviewee's were also enthusiastic about the use of an avatar or similar interactive agent that could 'guess' when you were having difficulties completing the form.

Finally, the sixth design requirement observed suggests that as with a paper-based form, whereby you can start a form and leave it to be completed at a later time or date, this too should be reflected in the functionality of an online version of the financial form and the housing form but not the welfare form. On reflection this may have been due to the length of the welfare form and the lack of additional/historical information (required for the financial and housing form) that the older adult may have had to seek out to complete the form, and thus did not require such a 'pause' in completion.

6 Conclusions and Future Research Directions

The access, completion and submission of online forms present a number of benefits to older adults, particularly reducing or eliminating difficulties relating to cognitive, sensory and physical declines observed as a product of the natural aging process. This paper has validated, to some extent, and extended earlier design requirements offered by older adults in relation to e-Government service access forms. Future work comprises the implementation of the design requirements offered here within a prototype system that will be employed as a system design evaluation tool.

If all requirements continue to be positively evaluated, future research may wish to address the impact that the online form design features have on each other in terms of redundancy. For example, if you have a help function and a pop-up box providing the older adult with assistance during form completion, will one feature counter the other?

Clearly greater research in this field is necessary if we are to promote a society that values and supports the independence and social inclusion of a rapidly aging population.

References

1. Barnett, R.: Forms for People. Robert Barnett & Ass. Pty Ltd. (1999)
2. Brouwer-Janse, M., Suri, J., Yawitz, M., De Vries, G., Fozard, J., Coleman, R.: User Interfaces for Young and Old, ACM Interactions 4(2), 42–46 (1997)
3. Government News Network. New Action to Tackle Social Exclusions and Link Services for Older People (2006) Available: http://www.gnn.gov.uk/content/detail.asp?NewsAreaID=2&ReleaseID=185469
4. Gregor, P., Newell, A.F., Zajicek, M.: Designing for Dynamic Diversity – interfaces for older people. In: Proceedings of ASSETS 2002, The Fifth International ACM Conference on Assistive Technologies, July 8-10, Edinburgh, UK pp. 151-1155 (2002)
5. Hawthorn, D.: Possible implications of aging for interface designers. Interacting with Computers 12, 507–528 (2000)
6. Hazlett, S.A., Hill, F.: E-government: the realities of using IT to transform the public sector. Managing Service Quality 13(6), 445–452 (2003)
7. Lines, L., Hone, K.S.: Eliciting user requirements with older adults: Lessons from the design of an interactive domestic alarm system. Universal Access in the Information Society 3(2), 141–148 (2004)
8. Lines, L., Patel, Y., Hone, K.S: Online form design: Older adults' access to housing and welfare services. In: Proceedings of HCI and the Older Population Workshop, Workshop Proceedings, HCI 2004, September 7-11, Leeds UK pp. 21–22 (2004)
9. National Audit Office, Improving & Reviewing Government Forms. Maxim: London (2003)
10. National Statistics. National Population Projections (2000) Available: http://www.statistics.gov.uk

An Accessible Platform for Conference Administration and Management

George Margetis[1], Stavroula Ntoa[1], Maria Bouhli[1], and Constantine Stephanidis[1,2]

[1] Foundation for Research and Technology – Hellas (FORTH),
Institute of Computer Science, Heraklion, GR-70013, Greece
[2] University of Crete, Department of Computer Science
{gmarget,stant,bouhli,cs}@ics.forth.gr

Abstract. Nowadays online conference management systems are an essential tool for the organization of a conference. However, most such systems do not provide support for accessibility by disabled users. This paper discusses the design and development of a platform supporting the development of accessible conference management systems, focusing on accessibility integration through the use of an iterative approach. Accessibility issues as they were raised during the evaluation procedure are also discussed.

Keywords: Accessibility, model-based web application development, conference management.

1 Introduction

Web-based systems for the management and administration of conferences have evolved into sophisticated tools for addressing the needs of conference participants and organizers, providing a large number of functionalities to support conference management tasks. There is a plethora of such systems available, commercial or open-source, supporting a variety of tasks, such as electronic submission, review, advocacy, and acceptance of papers for the organizing committee, or paper submission, editing and user profile facilities for simple users. However, support for accessibility is lacking or minimal in most cases.

On the other hand, the need for web resources to be accessible to people with disabilities has become compelling. A number of guidelines have been published by the World Wide Web Consortium, the European Commission, national governments and others during the last years in order to address the issue of web accessibility. The most well-known guidelines for web content developers are the Web Content Accessibility Guidelines (WCAG) [1].

However, creating an accessible web service is not just a matter of applying these guidelines. Instead, a more complex procedure is needed in order to ensure the desired level of accessibility, engaging web designers, web accessibility experts, web developers and of course representative end users. In this paper, the design and development of an accessible platform for conference administration and management is discussed, and the iterative procedure that was applied, as well as some of the

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 941–950, 2007.
© Springer-Verlag Berlin Heidelberg 2007

interesting accessibility issues that occurred, are presented. Accessibility evaluation was placed at the center of the iterative development procedure, resulting in minor redevelopments through the development lifecycle. Last but not least, it should be mentioned that the platform was designed to support not only the development of accessible conference management systems for users with disabilities, but also of system usable for everyone.

2 Background

Modern web based applications follow the Model-View-Controller design, where Model represents the application business logic or state (e.g., the data stored), while View represents the User Interface elements of the application and Controller represents the navigational code. In Figure 1, the lifecycle of a Model-View Controller design is represented, where the solid lines indicate a direct association and the dashed lines an indirect one. In more detail, the Controller is responsible for handling events (e.g., user actions) and updating the Model accordingly. Control is usually then forwarded back through the Controller to the appropriate View.

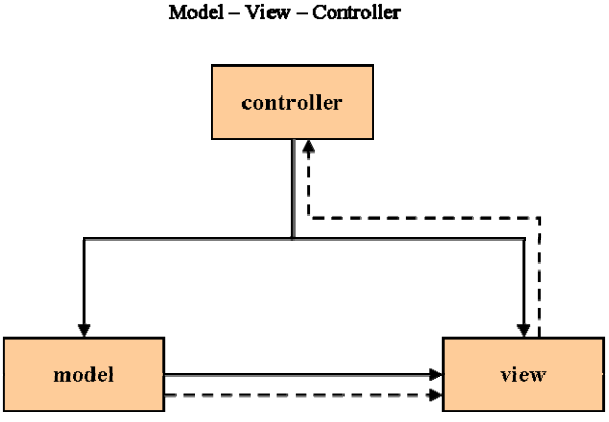


Fig. 1. Model-View-Controller design representation

The main advantage of this approach is that it allows the separation of the semantic information regarding the use of objects, their relations and meanings from the way this information is displayed to the users, providing thus the ability to isolate the user interface from the business logic of an application. Therefore, changes to the user interface do not impact the data handling, while data can be reorganized without changing the user interface. In modern web applications, data in the View part is displayed through a combination of templates, facilitating thus accessibility integration.

In order to facilitate accessibility in a database driven application, quality control on inserted data and to user interface must be ensured [2]. Quality control concerns the validation of new content for conformance with accessibility guidelines as a part

of the web application. Moreover, the evaluation process should include additional tests for the accessibility of content which is formulated as a result of the combination of templates encapsulating accessible content chunks. Interface accessibility refers to the employment of W3C Web Content Accessibility Guidelines to ensure the accessibility of form and navigation elements the user interacts with. In web applications that use templates as common containers for data representation, accessibility integration mainly results from accessible templates design. If templates are well-designed and the content creation process is strictly controlled for accessibility, the web application is likely to meet accessibility standards [3].

The purpose of the developed platform was to provide the necessary infrastructure in order to enable the development of accessible conference management applications. For the design and development of the platform, an open-source Content Management Platform [4] was used and appropriately expanded to incorporate all the functional modules needed to support the organization and administration of a conference. The open-source platform that was used enables the development of web applications following the Model-View-Controller design, which is suitable for accessibility integration. Moreover, it provides core functions for database communication, dynamic web content presentation with the use of templates and sub-templates placeholders, as well as users and permissions management.

Regarding the development of the accessible platform, quality control issues were negligible due to the nature of the content. More precisely, the content that is inserted in a Conference Management System mostly consists of unstructured plain text (e.g., paper title, author names, email addresses, dates, numbers, etc.) that can be easily encoded into a strict semantic scheme, such as XHTML, which is the necessary precondition for conformance with accessibility guidelines. Therefore, efforts towards accessibility integration mainly focused on user interaction and navigation issues, which will be further analyzed in the next sections.

3 Conference Management Platform Description

Besides support for accessibility, the main goal in the design and development of the conference management platform was to provide a variety of functionalities, in order to address the needs of both the conference organizing committee and of conference participants.

Functionality provided to conference participants includes:

- Paper submissions, in three stages: abstract submission, paper submission and submission of camera-ready version of the paper
- Tutorial and poster submissions
- Conference Registration
- User profile management.

Functionality provided to the conference organizing committee includes:

- Reviewing process administration, encompassing the assignment of papers to a number of reviewers, reviews administration and authors notification
- Program organization, including advance program and final program editing

- Submissions administration, allowing submissions categorization by thematic area as well as the administration of authors submissions
- Users administration
- Student volunteers administration
- Administration of participants' conference registrations.

4 Methodology

The process that was applied for the design and development of the conference management platform was iterative, mainly focusing on accessibility evaluation. Following the Model-View-Controller approach, the design of the conference management platform focused on creating accessible XHTML templates for every functional module. Therefore, the platform ensures that conference management systems based on it will have all the required functionality and will be inherently accessible. Conference management system developers will only have to adapt the look and feel of their application (e.g., style sheets, images, etc.).

The platform development process, as shown in Figure 2, was initiated with the Requirements Elicitation phase, the results of which were used for the design phase.

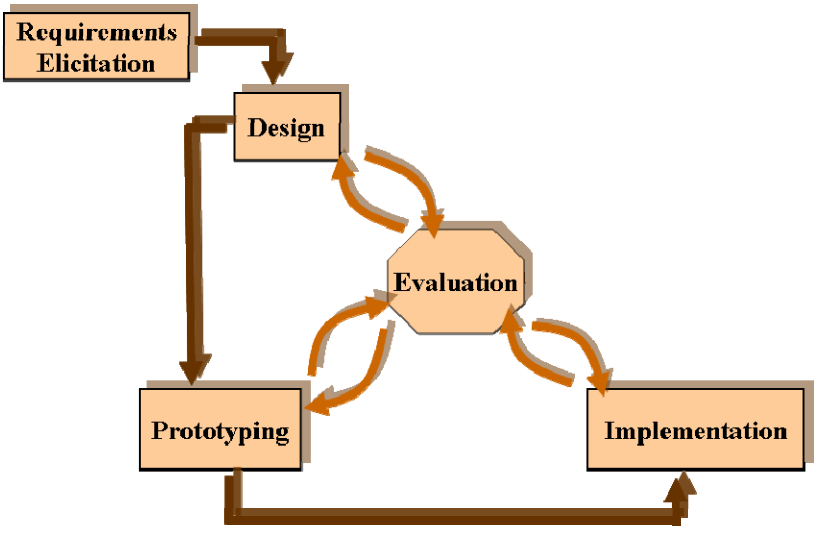


Fig. 2. Iterative Design Process

The design process included the creation of mock-ups and their evaluation by accessibility experts, using detailed checklists of accessibility features that should be later included during the development of the platform, based on W3C WCAG 1.0.

Once the design of the platform was finalized, interactive prototypes were created in XHTML. These prototypes were evaluated for their accessibility with the use of an automated tool [5].

Finally, after having ensured the accessibility of the prototypes, the implementation phase, based on these prototypes, was initiated. Implementation was alternated with accessibility evaluation until the desired level of accessibility was reached. A more detailed discussion regarding the evaluation of the implemented conference management platform is addressed in the section "Accessibility Evaluation".

After the prototyping phase, all the individual accessible templates that were required in order to compose the overall user interface were created. In particular, a main container template was used to structure the main user interface layout by defining five basic areas of functionality: header, navigation menu, navigation path, main page content and footer.

Header includes the application logo and the quick access links. Navigation menu comprises the main functionality provided, organized in a two-level menu. Menu options are differentiated according to the user role (e.g., conference participant, author, reviewer, administrator, etc.). Navigation path consists of a number of links of increasing granularity, according to the user selections. Footer displays information regarding the compliance of the application to the accessibility guidelines. Finally, the page content template can be further decomposed to a combination of containers of one or more templates, according to the specific system functionality. The position and style of the containers in which the sub-templates are displayed are specified by the page content template handler.

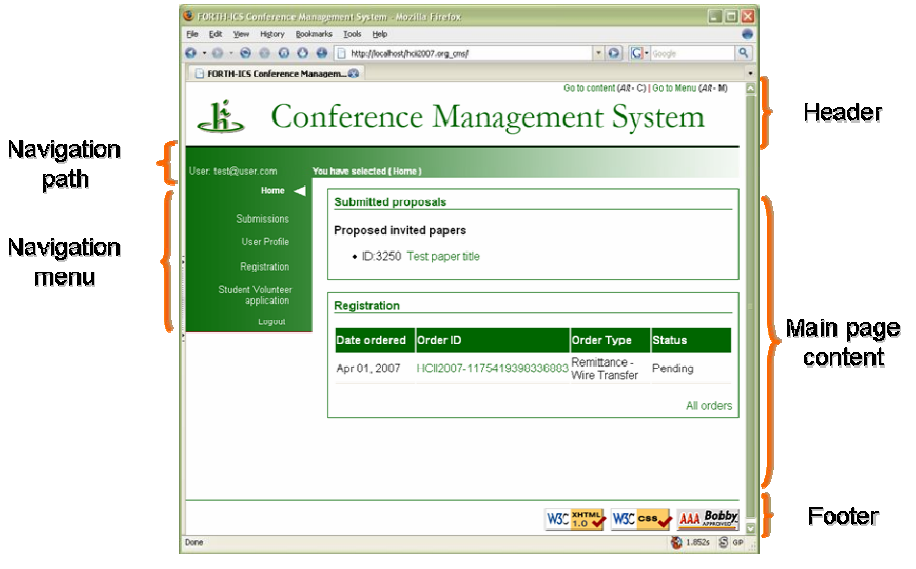


Fig. 3. Summary page of a user account explaining the use of five basic areas of a potential Conference management system

For example, Figure 3 illustrates the summary page of a user's account of a testbed conference management system, displaying information regarding submissions and registration, each constituting a different sub-template derived from a separate model handler.

5 Accessibility Evaluation

Accessibility evaluation during the implementation phase was carried out as a combination of manual evaluation and evaluation with the use of tools. A representative sample of pages was selected, including pages with a variety of layouts and functionality.

Manual evaluation of the selected sample included their examination with different graphical web browsers. The most popular web browsers were used (Internet Explorer, Mozilla Firefox, Opera) in their latest and one older version. For each browser, the following adjustments and checks were performed:

- Images were turned off and the availability and appropriateness of alternative text was examined
- Browser controls were used to change the font size, while page layout was checked to determine whether web pages were still readable at larger font sizes
- Different screen resolutions were tested and the page layout was checked to verify the absence of horizontal scrolling
- Color was changed to gray scale in order to determine whether the color contrast was adequate
- Keyboard was used to navigate through the page content, ensuring that access to all interactive elements was provided, in an appropriate order
- Scripts and style sheets were disabled and pages were tested to verify whether the functionality was still available and the layout was usable.

Furthermore, the selected pages were examined with the use of a text browser [6] in order to ensure that all the information and functionality were provided and that information was presented in a meaningful order.
In addition, the following automatic and semi-automatic evaluation tools were used:

- W3C Markup Validation Service [7], in order to check XHTML conformance to W3C Recommendations
- W3C CSS Validation Service [8], in order to check the style sheets that were used
- Watchfire® Bobby™ [5], in order to examine the pages' conformance with the W3C WCAG.

It should be mentioned that pages were also evaluated regarding the colors that were used, in order to verify whether the brightness and the contrast between the text and the background colors was adequate. To this end, two on-line free tools were employed. The first was AccessColor [9], a report generating tool, according to which both color difference and color brightness meet the recommended standards. The second tool was ColorFilter [10], simulating the look and feel of a page for various user categories with vision impairments, such as protanopia, deuteranopia, etc., according to which all the displayed information is readable by all the simulated user categories.

The result of these evaluations lead to some development-evaluation iterations, until the desired level of accessibility was reached. It should be noted that iterations were minor and the accessibility problems that were reported were few, mainly due to the accessibility evaluation procedures that were applied during the design and

prototyping phase. In more detail, during the design phase, accessibility experts provided guidance regarding major accessibility problems that should be avoided. Evaluation in the prototyping phase ensured the accessibility of each individual sub-template that would be used for the composition of the main interface. Therefore, accessibility evaluation of the implemented interface mainly focused on issues regarding the combination of the sub-templates as an overall accessible interface.

6 Accessibility Issues

During the iterative process of the design and development of the developed conference management platform, several accessibility issues were identified. Besides the well-known ones, such as the necessity of a descriptive alternative text for image, or the use of style sheets for the layout of a web page instead of layout tables, some other interesting concerns also came up.

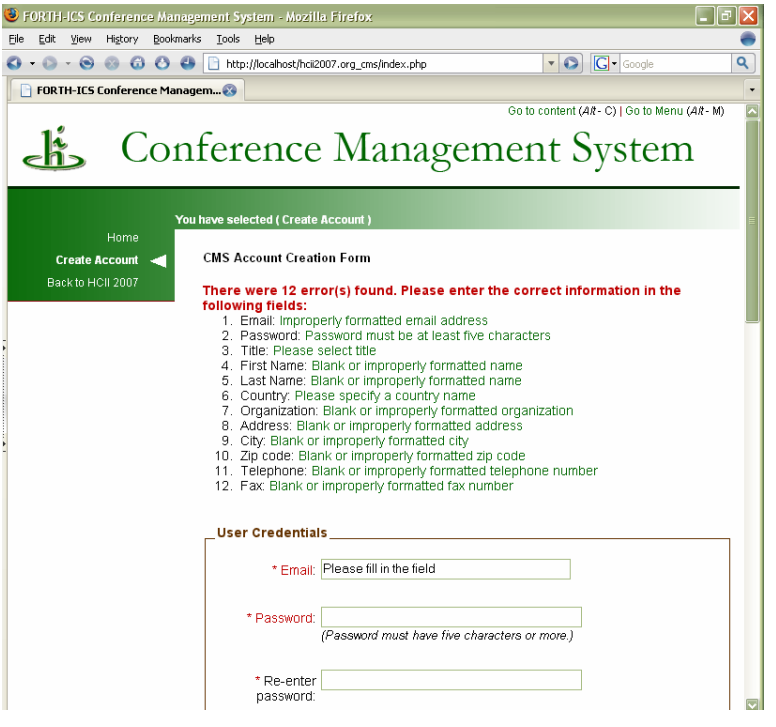


Fig. 4. Accessible error handling in forms

An interesting issue is that of error handling in forms. Usability experts usually suggest placing error messages in red color below the erroneously completed fields and to mark these fields with a red color as well. However, this approach is not usable by people with disabilities. In more details, the main drawbacks are the use of color to

signify information, which violates accessibility guidelines, and the increased effort required by blind users in order to locate the errors.

An alternative approach, as shown in Figure 4, is to place all the error messages to the top of the content, allowing direct access to error messages for users with vision impairments. Furthermore, error descriptions can be linked to the related field labels, enabling thus immediate interaction with these fields both for people with vision impairments as well as for people with motor-impairments of the upper limbs. Therefore, effective and efficient interaction is ensured in an accessible way.

Another issue worth mentioning is that of navigation. It should be possible for users of assistive technologies (e.g., screen readers) to easily and efficiently access the main page areas (e.g., menu bar, page content). This can be achieved with the use of quick access links on the one hand and skip links on the other. Quick access links are the topmost elements in every web page, providing fast access to main areas of the page and therefore the flexibility to directly move to a point of interest, avoiding interaction with preceding elements. For example, in the case of a blind user with a screen reader, interaction with the main page content is possible without the need of repeating already known information, such as the description of the site logo or the navigation path, on every new page request. Skip links in conjunction with quick access links further enhance user interaction, providing the ability to skip specific areas of information. These links are only visible when style sheets are disabled and page content is serialized.

An additional accessibility problem that occurred during the development was related to radio buttons that were initially selected to be used in forms. Some of the radio buttons accessibility drawbacks are:

- They do not scale along with the text on the page, making it difficult for users with low vision to determine whether they are selected and impeding users with motion impairments from selecting them because of the small target.
- Since in a radio buttons group many form controls are involved, explicitly associating the prompting text with form controls is impossible, misinforming thus users with screen readers.
- They introduce multiple form controls into the tab order, resulting in extra effort for screen reader users to navigate a form.

The alternative solution that was preferred was the use of selection boxes, which address all the aforementioned issues. In more detail:

- they scale along with the text
- the prompting text can be easily and accessibly associated to the select element since there is only one form control involved,
- due to the uniqueness of the form control, tab navigation is feasible.

Finally, the issue of handling deep link users [12] is also worth mentioning. Using bookmarks is convenient for people with disabilities, in order to directly access an often visited web page (e.g., information regarding an author's submissions review results). The developed platform supports stateless requests for most of the functionalities provided (i.e., users do not need to follow a specific interaction path), with the only precondition that user's browser has stored a unique identification cookie for this user, in order for the system to automatically log the user in.

7 Conclusions and Future Work

This paper has discussed the design and development of an accessible platform for conference management and administration. It was developed based on an open source Content Management Platform, enhancing accessibility with the use of templates. The process that was followed for the design and the development of the conference management platform was iterative, alternating the design, the prototyping and the development phase with accessibility evaluations. This approach resulted in minor redevelopments after each evaluation phase, reducing therefore the resources needed for accessibility integration.

This paper has discussed the main accessibility issues that were identified during evaluation, including error handling, navigation issues, radio buttons accessibility issues and handling deep link users. The aforementioned concerns are not only related to the conformance with guidelines, but also to improving the interaction for disabled users.

Future work will address issues of user interface adaptation to the users needs, providing for example alternative layouts for each user category (e.g., blind or motor-impaired users). A possible approach would be to use different sets of templates for individual cases, which is easy and feasible, since presentation and behavior are independent from the data structure.

Furthermore, future work will include user-based evaluation, in order to test the accessibility of the system with the users it addresses to and obtain metrics not only regarding its accessibility but also regarding the efficiency and effectiveness of users with disabilities with the system. Therefore, even though the system is accessible, several interaction enhancements may be also implemented to achieve shorter task completion times and higher user satisfaction rates.

A barrier to the development process was the suggestion of W3C WCAG 1.0 to avoid the use of scripts, since in certain cases using them was necessary in order to achieve higher usability. For example, until user agents handle empty controls correctly, default, place-holding characters should be included in edit boxes and text areas. However, the absence of javascript forces the user to select the place-holding text and delete it before entering text. This could be automatically handled with the use of javascript, by deleting the place-holding text whenever the field obtains the mouse or keyboard focus. In general, the use of javascript could improve the usability of the application, providing an alternative context organization and shorter task completion times. Recognizing the need for adaptation to new technologies, W3C proceeded to the study and development of a newer version of its guidelines in order to include the new trends in web applications [11]. Therefore, future work will address the issue of evolving the conference management platform according to an Asynchronous JavaScript and XML (AJAX) approach.

References

1. World Wide Web Consortium Web Content Accessibility Guidelines 1.0. http://www.w3.org/TR/WCAG10
2. May, M.: Integrating Accessibility into Content Management Systems. In: Proceedings of the XML Conference 2003, Pennsylvania, Philadelphia, USA (2003)

3. Burzagli, L., Billi, M., Gabbanini, F., Graziani, P., Palchetti, E.: The Use of Current Content Management Systems for Accessibility. In: Miesenberger, K., Klaus, J., Zagler, W., Burger, D. (eds.) ICCHP 2004. LNCS, vol. 3118, pp. 331–338. Springer, Heidelberg (2004)
4. Phpwebsite website http://phpwebsite.appstate.edu
5. Watchfire Bobby http://www.watchfire.com/products/webxm/bobby.aspx
6. Lynx text browser http://lynx.browser.org/
7. W3C Markup Validation Service http://validator.w3.org/
8. W3C CSS Validation Service http://jigsaw.w3.org/css-validator/
9. AccessColor - Online Tool for Colour Contrast http://www.accesskeys.org/tools/color-contrast.html
10. Colorblind Web Page Filter http://colorfilter.wickline.org/
11. Roadmap for Accessible Rich Internet Applications (WAI-ARIA Roadmap) http://www.w3.org/TR/aria-roadmap/
12. Nielsen, J., Loranger, H.: Prioritizing Web Usability, New Riders Press (2006) ISBN: 978-0321350312

Service Science: Applications of Observation to Real World Business

Haruhito Matsunami

Information & Communication Systems Dept., Osaka Gas Co.,Ltd
4-1-2, Hiranomachi, Chuo-ku, Osaka, 541-0046, Japan
matunami@osakagas.co.jp

Abstract. Original methodology of service science, that combines three methods (obtaining service data using ergonomics methods, analyzing service data using statistical methods, improving service using information technology) was constructed. Service science approach was implemented for interactions with customers such as daily life observation, shopping behavior observation, and talk analysis, and for employee's nonroutine tasks, such as gas pipe construction work observation. The service science methodology improves customer satisfaction and employee satisfaction by improving productivity and safety of the work.

Keywords: service science, observation, customer satisfaction, human behavior.

1 Introduction

Service science is a concept that applies scientific methods to service, which relied on intuitions and experiences, in order to increase the productivity and facilitate innovations. Service science is a developing new academic field, and the definition and methodology are still not established. But some companies started addressing the issues of service science to construct original methodologies for practical reasons.

In this paper, definitions, methodology, and examples of service science of Osaka Gas Co., Ltd, which is a gas supplying energy corporation, is discussed.

Service is becoming more and more important in worldwide economy. For example, more than half of the sales of IBM and GE are by service. In Japan, the ratio of service business in numbers of employees and gross domestic product is increasing. The importance of service is increasing as such, but scientific approaches, such as in manufacturing business, have not been taken in service business.

IBM Almaden Services Research suggested "service science". It is officially called "Service Sciences, Management and Engineering". There are varieties of definition for service, such as "the distinct characteristics of services are intangibility, perishability, heterogeneity of the product, and simultaneity of production and consumption" (Sasser et al. 1978)[1]. But the two important points in service science are as follow.

1) Service science applies scientific methods to service in order to improve.
2) Service science is an academic field that should be promoted by specialists in huge variety of academic fields.

R. Cooper and J. Kennedy (Eds.): BNCOD 2007, LNCS 4587, pp. 951–960, 2007.
© Springer-Verlag Berlin Heidelberg 2007

In the US, academic society is leading service science as a new academic field. In colleges in the US, departments of computer science are trying to build this new academic field service science with business schools, and some colleges have service science academic courses.

2 Approaches to Service Science

In this chapter, the methodology that Osaka Gas constructed is discussed.

2.1 Definition of Service Science

The general definitions of service are abstract, so we defined service as shown in Fig.1. The situations of service are also defined as shown in Fig.1.

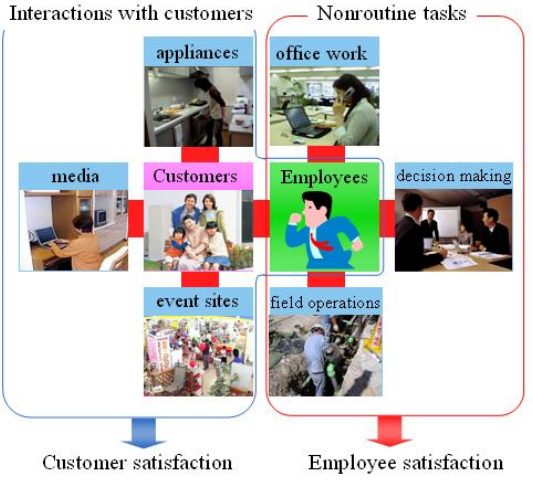


Fig. 1. Situations of service

The targets are customers and employees of the corporation. The situations of service for customers are four kinds of interactions between customers and corporation.

1. Direct interactions between customers and employees (face-to-face interaction at customer's house, call handling at call center)
2. Interactions at event sites organized by the corporation
3. Interactions through media (websites, TV, newspapers)
4. Home appliances that customers use

The situations of service for employees are three kinds of nonroutine tasks.

1. Office work with IT devices (computers, cellular phones, portal websites)
2. Decision making (formulations of effective strategies, contract negotiations)
3. Field operations (construction field, plant operation)

2.2 Definition of Service Science

In this study, service science is defined as "approach to apply three scientific methods for customer service and employee service to improve customer satisfaction and employee satisfaction" as shown in Fig.2. The three scientific methods are as follow.

1. Obtaining service data using ergonomics methods
2. Analyzing service data using statistical methods
3. Improving service using information technology

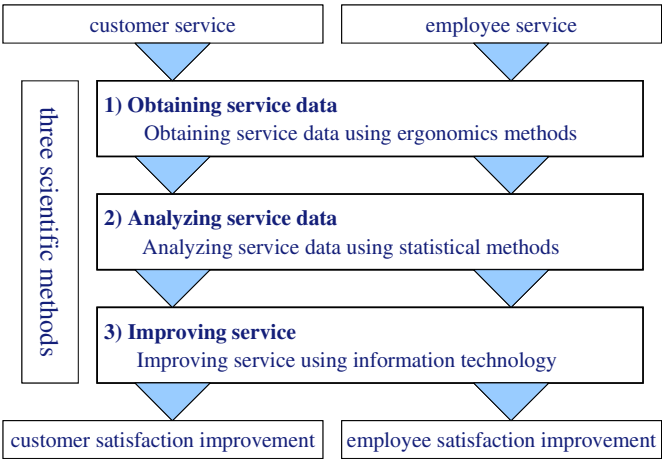


Fig. 2. Definition of service science

2.3 Methodology of Service Science

The data analyzed in service science is defined as service data. Service data can be divided into two categories; system service data and human service data. System service data consist of corporation operation system data (customer data and sales data), and available data from outside (market data and climate data).

Human service data consist of human behaviors, such as how customers use their home appliances at home, how they move in event sites, and how employees work in office and field. It is more difficult to collect and analyze human service data than system service data. In this section, three steps that constitutes of service science methodology are discussed.

2.3.1 Service Data Collection

In order to review service fundamentally, service field must be observed comprehensively[2]. Observation is a method to analyze human behaviors, performance, and circulation with specialized knowledge of ergonomics. Observation is often accompanied with interviews using psychology and ethnography skills. Compared to questionnaire, which is to investigate exposed needs and problems, observation is to find potential needs and problems. American advanced IT

corporations, such as IBM, Microsoft, and Intel, hired ethnographers and anthropologists, and investigated human behaviors at office and home through observation and interview, to develop new IT equipments and service. We also started collecting human service data by ergonomics methods as follow.

1) Observation
By observation method service fields are observed. Notes, that include words and figures, are taken for facts and findings. It is important to observe without individual prejudices because they can prevent from getting new findings. Fields are also videotaped when possible. Sounds and voices are sometimes recorded with IC recorder.

2) Interview
Interviews are performed to investigate true needs and values of the participants. Notes are taken and interviews are also videotaped when possible. It is important to build "rapport (individual relations of trust)" at first. Without rapport, it would be difficult to get profound insights through interview. In some interactions such as event sites, it is impossible to have interviews, so service data are taken only through observation in such cases.

2.3.2 Service Data Analysis

Human service data are analyzed as follow.

1) Obtaining findings
People in charge, such as ergonomists and sales people, group together and discuss on the video data and notes to obtain findings on human behaviors from the observation and interview. In case of 2-hour video about one hundred findings would be obtained.

2) Obtaining qualitative hypotheses
As the next step, interpretations are given to all the findings. Interpretation is a task to discuss potential needs and problems from findings and facts through ergonomics and psychology standpoints. Not all the findings can be the source of potential needs and problems, but these would be the qualitative hypotheses.

3) Examination of hypotheses by quantitative analysis
As the final step, the hypotheses are examined by quantitative analysis. Observation software "observer" was brought in to quantify human behaviors, performance, and circulation from video data. With this software, human behaviors, such as how many times participants did a specific task, how they moved, how long it took to do the task, can be quantified by watching the video.

2.3.3 Service Improvement

After the service data collection and analysis, IT solutions are provided to improve service as the third step.

For example, when technical know-how for event sites is obtained, database for the know-how is developed to share it among people in charge of events.

IT staff in agreement with needs of business staff has implemented system development. However, business staff does not always comprehend all the needs of the customers and the service field. The most significant characteristic of this

methodology is that service data are collected and analyzed, and then specialists in huge variety of academic fields improve the service.

3 Examples of Service Science for Customers

As shown in the Fig. 3, the situations of service for customers are four kinds of interactions between customers and corporation, such as 1) direct interactions between customers and employees, 2) interactions at event sites organized by the corporation, 3) interactions through media, and 4) home appliances that customers use. In this chapter, the examples of these four interactions are discussed.

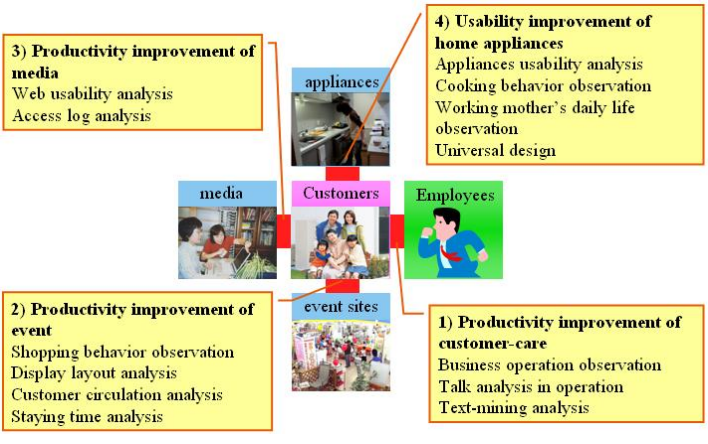


Fig. 3. Four kinds of interactions between customers and corporation

3.1 Direct Interactions Between Customers and Employees

Interactions between customers and employees include face-to-face interaction at customer's house and call handling at call center. An example of interaction at the visit to customer's house is discussed. Before household customers begin to use gas after move, worker visits customer's house to check gas pipes to start using gas. At the occasion, worker explains basic precautions.

1) Service data collection

As a part of customer satisfaction investigation, worker's attitudes and understandability of explanation in the visit were investigated for randomly sampled customers to find that there were significant differences between advanced-level workers and intermediate-level workers.

In order to clarify the factors of the differences between the advanced and the intermediate, researchers are accompanied with the workers to observe their activities and record their talks.

2) Service data analysis
Collected voice data were converted into text data and analyzed by text-mining approach. As an example of results, intermediate-level workers used foreign words significantly more often than advanced-level workers. In the explanation of appliances foreign words are likely to increase by 1.5 times, and in the explanation of inspection they are likely to increase by about 3 times. The foreign words in the Japanese explanations are English words such as “adaptor”, “full name”, “controller”, “check”, and “cap”. Especially for elderly customers, using foreign words is one of the factors that make the explanations difficult to understand.

3) Service improvement
Workers were educated not to use foreign words in the explanations.

3.2 Interactions at Event Sites Organized by the Corporation

Osaka Gas organizes many events to invite customers. An example to improve events is discussed.

1) Service data collection
In a typical event titled “gas-ten”, video cameras were installed to observe human behaviors at each zone as shown in Fig.4. During the 4-day event period, about 120-hour video was recorded. In other event, another 120-hour video was recorded, so 240-hour video was obtained in total.

Fig. 4. Video at event site

2) Service data analysis
As an example of analysis, customer’s circulation in the event site is discussed. An event site was round, and the display layout was designed so that the customers would move in the anti-clockwise direction from the entrance. However, on the 1st day, the video showed that the circulation was clockwise for 60% of the customers.

3) Service improvement
After the event on the 1st day, the display layout was improved so that the circulation would be anti-clockwise naturally. On the 2nd day, the circulation became anti-clockwise for 80% of the customers as expected. Other than this improvement, know-how, such as poster display method to extend the customer’s stay and product display

method to increase customer's browse. The database for obtained technical know-how for event sites was developed to share it among people in charge of events.

3.3 Home Appliances That Customers Use

In this section, the example of product development is discussed.

1) Service data collection
55 households of customers were visited to observe human behaviors in kitchen, bathroom, and living room. The behaviors were videotaped, and 170-hour video was obtained as shown in Fig.5.

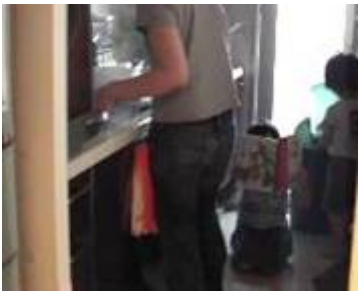


Fig. 5. Human behavior data in kitchen

In order to comprehend the current living conditions, cooking behaviors for dinner of each generation were videotaped. Working mother's activities for a whole day were also observed to know their current living conditions. In addition to in-door housework, such as cleaning, laundry, and communication with children, out-door behaviors, such as shopping, pickup and drop-off, were also observed.

Both in in-door and out-door, the observation notes were taken, and the in-door behaviors were videotaped. After or during the observation, interviews are performed to investigate true needs and values of the participants.

2) Service data analysis
As an example of the analysis, development of new gas cooking stove is discussed. By analyzing the cooking behaviors, about 800 findings were obtained. 20 important findings were selected by being evaluated through the technical viability.

3) Service improvement
Design proposal of new gas cooking stove was developed based on important findings. Group interview evaluation of the findings was held with housewives, and the final prototype was completed.

The findings obtained in this observation were not exposed needs, which were available through conventional questionnaires, but potential needs. Therefore the findings are gold mine of ideas for new gas appliances. Video data are also valuable because they can be basic data for variety of purposes. For example, the video of cooking behaviors can be applied not only to product development, such as gas

cooking stoves, dishwashers, and system kitchen, but also to kitchen cooking devices, shape of food, package improvement.

4 Examples of Service Science for Employees

As shown in the Fig. 6, the situations of service for employees are three kinds of nonroutine tasks, such as 1) office work with IT devices, 2) decision making, and 3) field operations. In this chapter, an example of these three interactions is discussed.

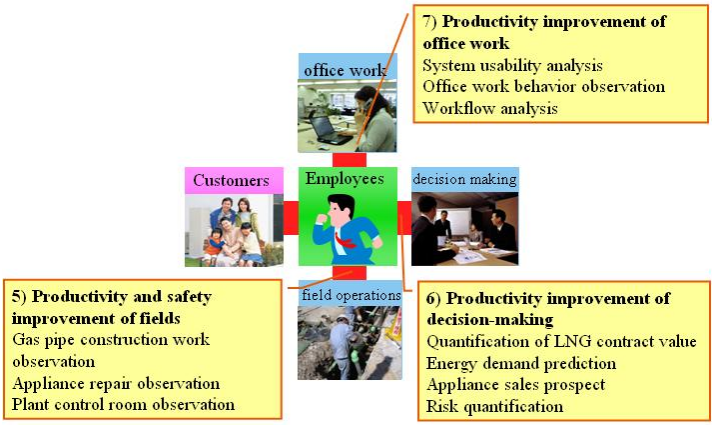


Fig. 6. Three kinds of nonroutine tasks

4.1 Field Operations (Construction Field, Plant Operation)

Osaka Gas has many construction fields such as maintenance work of gas facilities in customer's house, new constructions of gas conduit, and operation of gas plants.

Service to improve safety and productivity of field operations was addressed. In this section approach to improve safety of gas conduit is discussed.

1) Service data collection

Osaka Gas is proceeding activity to reduce the number of accidents in construction work of gas conduit into zero. We have had schemes to educate workers based on the preventative measures of actual accidents. Now it is aimed to introduce security measures and prevent accidents not only from "exposed" error information but also by obtaining "potential" error risk factors from the field.

Ergonomist is accompanied with the workers in the field to observe their activities as shown in Fig. 7. Activities in the field were videotaped, and workers were interviewed. Two cameras were used; one for the whole field and the other one for the individual operation. About 100-hour video and 400 notes have been obtained.

Fig. 7. Gas pipe construction field

2) Service data analysis
100 potential risk factors that could lead to human errors were obtained by video data analysis. Interpretations were given to all the risk factors from ergonomics standpoints.

3) Service improvement
For all the 100 risk factors, solutions were suggested and database for the know-how was developed to share it among people in charge of construction work.

Using the same video of the construction fields, improvement of productivity is also addressed. For example, by measuring time for each step in the construction fields, chart analysis was implemented.

5 Conclusion

In this chapter the approach of service science in Osaka Gas is summarized and future plans are discussed.

1) Osaka Gas's original methodology of service science was constructed.
"Service" and "service science" in Osaka Gas were defined.

The service science methodology that combines three methods (obtaining service data using ergonomics methods, analyzing service data using statistical methods, improving service using information technology) is developed.

There are very few examples to suggest service science methodology in industrial world. Therefore this methodology will be referred when service science is addressed in other corporations.

2) Service science approach was implemented in the interactions between customers and Osaka Gas.
In gas pipes checking tasks, education manual to explain in an easily understood manner was developed.

In event sites, customer's circulation percentage in the intended manner was increased from 40% to 80%.

About 240-hour event video and 170-hour daily life video were obtained.

3) Service science approach was implemented in nonroutine tasks.
Construction fields of gas conduit were investigated to obtain 100 potential risk factors and make manual.

About 100-hour video of gas conduit construction fields was obtained.

Service science approach has started in Osaka Gas to improve service to high level, but it still has many challenges.

1) Expansion of application range
Targets of service science up to the present are only a part of the fields. We still have many fields that have not been investigated, such as office work, call center work, and service shops. Application range of service science should be expanded.

2) Improvement of methodology
New technology evolution on observation, analysis, and IT, will be necessary to expand the application range. For example, service science approach to office work requires communication analysis technology. Video has been analyzed by hand, but image processing technology will be necessary to quantify human behaviors. Methodology should be evolved by expanding alliance with other organizations.

3) Total management of service
Total management system of service in Osaka Gas should be addressed to integrate many different kinds of service.

References

1. Sasser, W.E, Olsen, R.P., Wyckoff, D.D.: Management of Services Operations, Massachusetts: Allyn & Bacon (1978)
2. Kelley, T.: The Ten Face of Innovation: Currency / Doubleday (2005)

VoiceBlog for Blind and Weak-Eyed People

Yoshie Sagata, Masahiro Watanabe, and Yoko Asano

Cyber Solutions Laboratories, Nippon Telegraph and Telephone Corporation,
1-1 Hikari-no-oka, Yokosuka-shi, Kanagawa, Japan
{sagata.yoshie,watanabe.masahiro,asano.yoko}@lab.ntt.co.jp

Abstract. We introduce VoiceBlog, a blog reading browser based on the concept of universal design. Existing screen reading software used by blind people for PC access fails to support rapid understanding of the original blog structure and contents. VoiceBlog has a cascading user interface, hierarchically structured, and can well parse strictly-structured Web content such as blogs. It uses an audio user interface rule, which simplifies the comprehension of contents and menus (arrangement of contents based on construction, simplification, and navigation) to allow blind and weak-eyed people to comfortably access blogs. Blind people can use the keyboard to access all functions. The arrangement and design of icons, colors and correspondence to hi-contrast functions can be customized for the weak-eyed and color blind. Four blind and four weak-eyed people participated in trials of VoiceBlog. The results showed that it allows the blind to access blog contents easily because keyboard manipulations are similar to usual reading browsers and screen readers. The shapes and color of the icons are easy to recognize for weak-eyed people.

Keywords: Universal design, blind people, blog reading browser, blog, hierarchic structure, cascading user interface.

1 Introduction

Universal design (UD) of information is important, and various studies have been made in response to the aging of society. UD of the Web should ensure that everyone, including senior person and persons with disabilities, can access information on the Web[1][2][3]. NTT has been studying universal design technology[4][5][6]. Recently, the use of Blogs, by which the individual can easily disseminate information over the Web, has increased rapidly. However, screen readers and voice browsers (hereafter, reading out software) fail to permit easy understanding of the content and structure of Blogs. We developed the approach of analyzing the layered structure of Blogs and used it to develop the reading out software VoiceBlog[7]. This paper reports on the interface design of VoiceBlog as it impacts the blind and weak-eyed people.

2 Problems with Screen Readers

The people who use voice reading out software include many with weak-eyes. This is because it is easier to hear the text spoken that to try to read it. In many instances,

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 961–969, 2007.
© Springer-Verlag Berlin Heidelberg 2007

such users often try to follow the text while listening to the speech. Therefore, it is necessary to optimize the display arrangement in addition to the reading software. The following problems exist when the blind and the weak-eyed read blog articles with existing reading out software.

First of all, when a blog is read out, it is sequentially read from the top of the page. In many cases, the menu of the blog and the title list corresponding to the menu are read early on, and articles cannot be reached easily. In addition, it is difficult to understand the blog's structure.

If the zoom function is activated, mouse movements become constrained to the displayed area[8]. This makes it very difficult to move between the top and bottom of the original page.

In addition, many weak-eyed users employ the high contrast mode of Windows Accessibility Option because it prevents screen reflections from preventing character from being seen[8]. The high contrast mode changes the background color and the character color to those set by the user. In general, the background color is black and the text colors are white/yellow/green. Another problem is that if high contrast setting is selected, some screen/menu elements become very difficult to discern due to poor or non-existent contrast.

3 Overview of the System

Blind users stated they desire for an Auditory User Interface (AUI) that specializes in aural presentation. Additionally, for the weak-eyed, icons were centralized to reduce the distance over which the mouse had to be moved. The icons were modified so that they could be easily read even if the high contrast function was used. Color combinations were developed that perceived even by those with color vision troubles.

3.1 Cascading User Interface

A key feature of blogs is their complexity; they offer so many different elements. The sighted user has complete freedom to scan the page and determine how he will access the contents. It is possible to reach the blog article desired from among the displayed contents by one click operation. However, because the other users must sequentially listen to all information, it is difficult to understand the blog's structure (links to articles, menu and so on). Therefore, it is important to create a process the handicapped user is presented with just that information needed to make the next control operation. We decided to convert the blog structure into a cascading display of menus and articles; the intent is to simplify the acquisition of information and subsequent navigation actions.

Figure 1 shows the system structure of VoiceBlog. In the Display/Operation part, VoiceBlog sends an http request via Internet to the goo-blog server after the user selects the title of a blog or inputs a URL in the Display/Operation part. Upon receiving the desired HTML source file from the server, VoiceBlog analyzes the file and determines which texts should be read aloud. The reading part analysis engine uses selection rules based on the special tags and the structure of the HTML source. The reading part analysis engine sends the texts to the synthesis engine which

verbalizes the texts through the PC's speaker. At the same time, the texts as well as the images, if any, are displayed on the PC's screen by the Display/Operation part. The voice synthesis technology used in the synthesis engine is "Fluet", which is based on phonological knowledge and clustering[9]. "Cralinet" which is a text-to-speech synthesis technology based on the corpus-based approach can also be used[10]. The Fluet version of VoiceBlog was used in the below analyses.

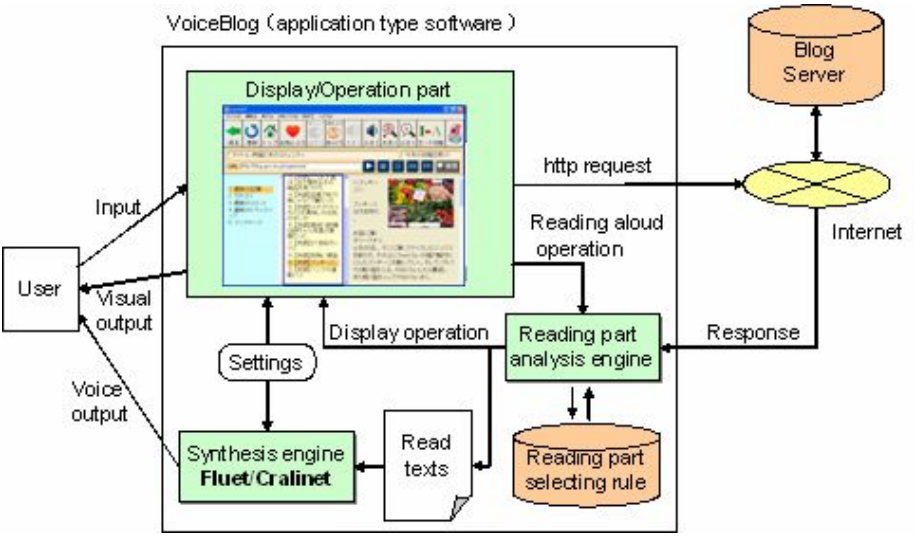


Fig. 1. System structure of VoiceBlog

The top page of goo-blog[11], a very popular commercial web site, is extremely busy with many menus, lists of contents, login panels, various templates, "New article", and "Access ranking" and so on. However, it takes too long to reach a target article if the article lies at the bottom of the page. VoiceBlog optimizes the content selection process and content presentation order to significantly improve blog article access. That is, the top page of goo-blog is replaced with a single menu whose contents and content order are decided by daily use. From an interview of 16 blog users, we decided to use the order of "1.New article", "2.Access ranking", "3.Pick up blog" and "4.Official blog". We added "5. Search" to these four entries. For blog pages we decided to set five entries on the top page: "1.Recent article", "2. Category", "3.Recent comment", "4.Recent track back", and "5.Bookmark".

3.1.1 VoiceBlog Graphics Components

Figures 2 and 3 show the graphic components of VoiceBlog. There are nine graphic elements.

1. Blog Operating Icon (11 icons : back, volume, text size etc
2. System State It shows system state, accessed blog information etc.
3. Blog Title Blog Title user selected
4. Post Number (The number of author posted article

5. Text Box for URL display and input Blog URL
6. Sound Controller 6 icons: for reading and finding keyword
7. Menu Area The goo-blog Selecting menu and blog operating menu
8. Title Area Blog title lists and blog article title lists
9. Content Area Blog article selected blog title

Fig. 2. Example of the goo-blog top page via VoiceBlog

Fig. 3. Explanation of VoiceBlog graphic components

The top screen of VoiceBlog shows two sub-windows; menu on the left and title list on the right as shown in Figure 2. Each blog screen accessed by a link from the title list on the top screen shows three sub-windows; menu on the left side, title lists in the center side, and target article on the right. Only the information needed to make a selection is read out sequentially because the menu and the title list are delimited by the cascading structure. This makes it easier and more efficient to reach the blog article text; only three short operations are needed (Operation1:Blog menu selection, Operation 2:Blog title selection, Operation 3:Reading of target blog article).

3.2 User Interface for Blind and Weak-Eyed People

When designing the User Interface (UI) we paid particular care to the needs of blind and weak-eyed people as described below.

3.2.1 Universal Access

The blind operate the software only through the keyboard. In VoiceBlog, all the operations usually done by mouse operation can be also done by keyboard operation. In keyboard operation, menu movements are made via the Tab key and upper and lower, right and left keys. The shortcut key was set to the main function. The key allocation of the shortcut key was set to duplicate, as much as possible, the same allocation offered by Internet Explorer (IE), the screen reader, and a voice browser, and so on. As a result, current operation experience is most effectively reuse. For instance, the page can be updated with the CONTROL key or the F5 key.

3.2.2 Layout and Design of Icons

When the zooming function is used by weak-eyed people, the screen area available to support mouse operation is very limited. Various operation icons and sound control icons were placed in the upper part of the screen to ensure that as many icons as possible remained in view and available for selection. Various operation icons were made much larger than IE icons to make mouse operations easier.

3.2.3 Screen Color Selection

People with visual disabilities often experience difficulty with different color combinations and they have their own preferences. To accommodate this need, the display color of the screen should changed to suit the user. Four color combinations (Type 1-4) that consider even the color blind were developed by using the color vision simulation software Vischeck[12]. Type 1, for seniors, uses a white background color black text, see Figure 4. Type 2, for the weak-eyed, uses a black background and white text, see Figure 5. Types 3 and 4, shown in Figure 3, are well preceived by the color blind.

3.2.4 Supporting the High Contrast Function of Windows Accessibility Option User

When the high contrast function of Windows Accessibility Option is used, the black background and white text of the standard PC display are reversed and most areas of

Fig. 4. Example of type 1 screen (background color: white, text color: black)

Fig. 5. Example of type 2 screen with hi-contrast function active (background color :black, text color:white)

the screen background are displayed in black. It is necessary to design the menus so the icons and other control buttons remain visible. Figure 6 shows an example of the screen when type 2 is selected and then the high contrast function. When the gray background areas of various operation icons are changed to black, it often becomes difficult to see icons and text, see Figure 6(a). the solution is to use white lines to delineate the icons and text, see Figure 6(b).

Fig. 6. (a) Upper: Icon and text contours without white border (b) Lower: Icon and text contours with white border

4 Evaluation by Blind and Weak-Eyed People

We investigated VoiceBlog in a trial by five blind and three weak eyed people. The execution period was about one week. At the beginning of the trial, all subjects were given a CD-ROM and told to install VoiceBlog on their personal computer and use it as their default browser for one week. All subjects were registered to the same mailing list through which they received daily questions (via e-mail) about their experience with the installation process and the Help file of VoiceBlog, what were their preferred functions, and the operation of VoiceBlog overall. We provided guidance via e-mail when the subject posed some question about operation. The

Table 1. Evaluation Results

Category	Subject	Comments made via e-mail
Structure	Blind	Can access article text without an extra reading out.
		The layout of the screen is comprehensible.
Sound	Blind	The tone quality of the effect sound firmly felt; jarring.
		It is not easy to catch it because it differs from usual screen reader; reading out tone quality and reading out method are different from normal.
	Weak-eyed	The effect sound while operating is useful.
	Both	The volume of the effect sound is larger than that of reading out.
UI	Blind	The key operation corresponds to known key operation.
	Weak-eyed	The shortcut key is enhanced.
		It is easy to see blog operating icons and the sound controllers.
		Even if the high contrast function is used, the icons can be recognized.

comments of the subjects are summarized in Table 1. The demands from the subjects are summarized in Table 2.

Table 2. Requests for VoiceBlog enhancement

Category	Subject	Comments made via e-mail
Guidance	Blind	Voice guidance function concerning operation
	Weak-eyed	Balloon function to recognize function of icons
Install	Both	The installation begins automatically.
	Blind	Repeat explanation of VoiceBlog
Help file	Blind	Easy explanation of various terms even if it hears
	Weak-eyed	The Help file can be displayed in the user's color setting.
	Both	Explanation of key operation with use
Blog article operation	Both	Copy function addition of blog title and content of article
		Link function addition from blog article to other Web sites
		Comment function and track back function addition
Color	Weak-eyed	Color customization to be user set

5 Discussion

Our study considered two VoiceBlog features; improving the efficiency of accessing blog articles by the cascading blog structure and UI that is effective for both blind and weak-eyed people.

Table 1 shows that blind people are comfortable using known key operations while weak-eyed people can see display items clearly even if the high contrast mode is used. Table 2 shows the weak-eyed subjects asked for control over the display colors and both groups wanted various voice guidance functions.

It is necessary to develop some intuitive function that allows the menu window items to be increased or reduced. For example, some subjects asked that the blog menu window be extended with the addition of comment and trackback posting functions.

6 Conclusion and Future Work

Blog access is becoming more frequent. However, the complex structure of blogs is not well handled by screen readers or voice browsers. This article introduced VoiceBlog; it offers several features to allow the blind and weak-eyed to access blogs much more easily and efficiently.

VoiceBlog has two main features.

- Its hierarchical display, based on blog structure, yields efficient access to blog articles
- Its UI is highly supportive of the visually disadvantaged Universal access, layout and design of icon, Color selection, adapted for high contrast function

Trials showed that it allows blind and weak-eyed people to well understand the structure of a blog. VoiceBlog also targets those normal users who have become tired with reading screen after screen; they are likely to find the synthesized voice highly

interesting. Computer novices are another likely audience. We intend to develop informative pop up labels for icons that provide various voice guidance instructions.

References

1. Web Content Accessibility Guidelines 1.0 http://www.w3.org/TR/WCAG10/
2. Japanese Industrial Standards Committee, JIS X-8341-3 Guidelines for older persons and persons with disabilities – information and communications equipment, software and services – Part3 Web Content: (in Japanese) (2004)
3. Universal IT Design Center, NTT http://www.waza.jp/ud_e/index.html
4. Asano, Y., Watanabe, M., Hamano, T., Ogawa, K.: Web Accessibility Guidelines that Solve Problems Peculiar to Japanese. In: Proceedings of WWCS 2004, pp. 601–606 (2004)
5. Watanabe, M., Okano, A., Asano, Y.: Universal Design Guidelines for Web Contents. NTT Technical Review 3(11), 17–22 (2005)
6. Watanabe, M., Okano, A., Asano, Y., Ogawa, K.: VoiceBlog: Universally Designed Voice Browser: IJCHI (In press)
7. Asakawa, C., Itoh, T.: User Interface of a Home Page Reader. In: Proc. the 3rd International ACM SIGCAPH conference on Assistive Technologies(ASSETS'98) pp. 149–156 (1998)
8. Make Windows XP More Accessible for Everyone, Microsoft http://www.microsoft.com/windowsxp/using/accessibility/default.mspx
9. Yoshida, Y., Nakajima, S., Hakoda, K., Hirokawa, T.: A new method of generating speech synthesis units based on phonological knowledge and clustering technique. In: Proceedings of ICSLP '96 pp. 1712–1715 (1996)
10. Mizuno, H., Asano, H., Isogai, M., Hasebe, M., Abe, M.: Text-to-Speech Synthesis Technology Using Corpus-Based Approach. NTT Technical Review 2(3), 70–75 (2004)
11. goo-Blog, NTT Resonant Inc. http://blog.goo.ne.jp/
12. Vischeck http://vischeck.com/

Inflection of In-Situ Ergonomics in Service Engineering

Atsushi Shinya

Systems Business Unit. Shimadzu Business Systems Corporation,
1, Nishinokyo-kuwabaracho, Nakagyo-ku, Kyoto 604-8442 Japan
shiny@shimadzu.co.jp

Abstract. "In-Situ Ergonomics" which extended the direct observation method was discussed. It is the method of using a sensor for data collection. By observing all the situations in which a user works for a long period of time, this method can be used in order to improve service. As an example, the use situation of the electric pot with a communication function was analyzed. Various users' use situation over one electric pot was clarified. And the improvement proposal of service was shown.

Keywords: In-Situ Ergonomics, direct observation method, sensors, datamining.

1 Introduction

Process P and Process D in a PDCA (Plan - Do - Check - Action) cycle have so far been well examined by service engineering. From now, the study of Process C and Process A will become more important. The Process C and Process A are input data of the next cycle of PDCA cycle(Fig. 1).

Quality of output cannot exceed quality of input. This paper describes research of the improvement method of Process C and process A.

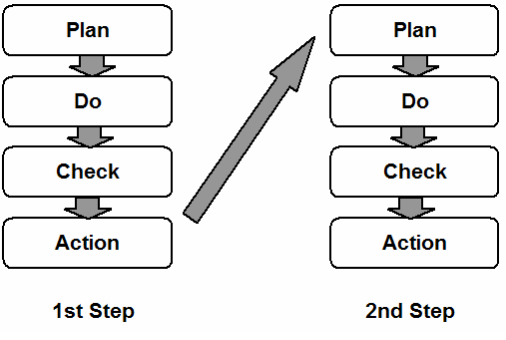


Fig. 1. PDCA Cycle

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 970–978, 2007.
© Springer-Verlag Berlin Heidelberg 2007

2 Direct Observation Method

The direct observation method is an objective method of collecting data.

The direct observation method involves the investigator witnessing the user in the environment where the product is used and recording the use over several hours or several days. This method has the benefit of being cheap, as it does not require especially high plant and equipment costs.

Questionnaires are a well-known method of surveying users. As a questionnaire poses questions set by the investigator, it can only obtain replies within the scope hypothesized by the investigator. Questionnaires are a method of extracting problems from the viewpoint of the investigator.

On the other hand, the direct observation method observes the user's actions to discover any problem points. It can therefore obtain results more closely matching the user, such as discovering unexpected ways of operation the user may adopt.

The direct observation method allows the investigator to discover and compile information on a user's subconscious behaviors. Consequently, it allows discovery of service problems that the user does not sense as a problem.

The direct observation method can obtain more profound user information but it is not without its shortfalls.

The first is the high personnel costs. The investigator stays in the subject's house and captures information on the user's behavior on a one-to-one basis. This contrasts with questionnaires that work on a one-to-many basis and simply require the replies to be recovered. The amount of data compiled in unit time is significantly higher using a questionnaire.

As the user controls the operations with the direct observation method, there is no guarantee that the user will use the service anticipated by the investigator during the period of the investigation. Even if the service is used, it cannot be guaranteed that this covers the entire usage pattern for that service by the user.

Also, with observations of just a single user, it is impossible to generalize the problem points with the service. As number of samples is small, the proportion of all users having that problem is unknown. Consequently, it is not possible to sort multiple problems in order of frequency of occurrence at the planning stage to determine countermeasures.

If the investigator discovers an unexpected problem through observations of a single user, there is a good possibility that the service has other problems.

The author has proposed In-Situ Ergonomics to compensate for the shortfalls of the direct observation method.

3 In-Situ Ergonomics

The ergonomic study domain apply in a PDCA cycle of service. In-Vitro Ergonomics will be applied in Process P and Process D. And In-Situ Ergonomics will be applied in Process C and Process A.

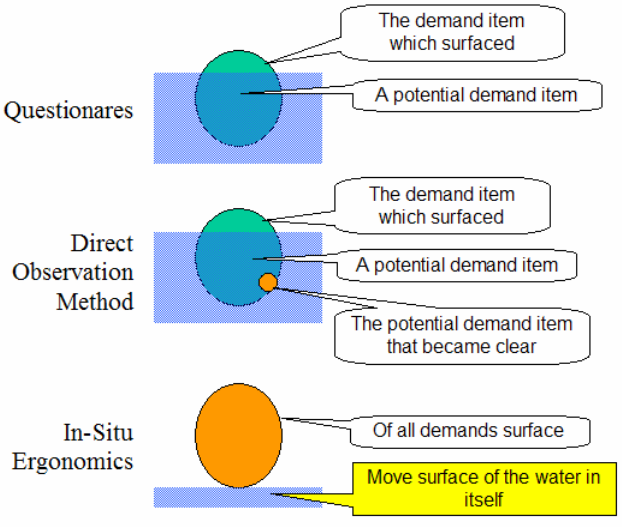


Fig. 2. The Iceberg Problem

Until now, the ergonomic study has been performed between Process P and Process D. In the beta test phase like Process P and Process D, a researcher researches for improvement of services.

Technology evolves now.

Various sensors can collect a product and information of the use spot of service. And information can transmit a message in the company side via a network. It is used for remote monitor of a control unit. And it is used as monitor information of a senior person, too.

For example,

- opening and shutting information of a door with microswitch
- On/Off of illumination using a pyro-electric sensor
- consumption information of water of a restroom with an automatic faucet.

The many information is provided. An ergonomic method called as "In-Situ Ergonomics", it collects and analyses data from a provided sensor in a product and service on the site. In other words, it becomes direct observation method with a sensor.

In-Situ Ergonomics solves "The Iceberg Problem" (Fig. 2).

By the direct observation method, a researcher and a user investigate it by 1 to 1.

By In-Situ Ergonomics method, it is advantageous in that following

- It can collect mass information at a time. It collects the information with a sensor containing it in product oneself via a network. This method can collect information of all the products in the using spot.

– In-Situ Ergonomics can conduct a continuous investigation. Whenever a user uses a product, information is dispatched automatically by a product installed in a product use place. On this account this method can investigate an action of a user continuously.

By these advantages,

– The researcher does not participate in data collection
– Large human reduction in cost is possible
– This method does not perform artificial data collection, it can collect an objective fact than conventional ergonomic study methods.

However, there are the following weak points in this method.

– Initial cost is big. Infrastructure maintenance such as in-home networks is necessary.

In-Situ Ergonomics applies service, and it is the method that is necessary to be improved.

4 The Service Engineering Application Example That Using In-Situ Ergonomics Method

The electric pot which has communication function were used. The use situation of 3,000 electric pots was investigated. A period is one month. Based on provided information, "The service that is serving hot water " is researched.

This service is called "i-Pot" and it is provided by ZOJIRUSHI. It is the service which watches elderly people from a remote place.

Fig. 3. The "i-Pot" service. (www.mimamori.net).

Each event of the "power supply ON (ps)" of an electric pot, "hot-water supply (hw)", and a "keeping-warm state (kw)" is acquired, and it transmits to a server through a communication line in real time.

A "keeping-warm state" is an event automatically transmitted for every hour, when an electric pot is in a turning-on-electricity state. A server sends the event information on an electric pot by E-mail according to a demand of the family of a remote place. Thereby, it enables the family of a remote place to watch elderly people's condition nonchalantly.

With this service, a user acquires the event of the electric pot currently used usually. For this reason, it is an advantage that a mental burden is not placed on the user side. However, the user using an electric pot is not necessarily elderly people. Moreover, it cannot be concluded that one person uses.

In other words, all the events of all the users using the electric pot can be acquired.

4.1 Data Specification

Although an electric pot transmits event information to a server in real time, the information summarized per 10 minutes here is used. This is equivalent to performing the time specimen method in a direct observation method in observation unit 10 minutes. The priorities of an event are 1. hot-water supply, 2. power supply ON, and 3. keeping warm.

It is represented with hot-water supply event when hot-water supply event and the power supply ON event occur in observation unit time.

The communication line is using "DoPa" of NTT DoCoMo.

4.2 Survey Content

It was observed by the In-Situ Ergonomics method for one month from November 1, 2006 to November 30. The number of object electricity pots was about 3,000 sets.

However, since there were an addition of a new contract person, a unused day, etc., the number of operation of day by day was changed.

Moreover, a user's attribute is unknown entirely. It is not used this time.

The calendar of Japan in November, 2006 is shown below. * The mark is a public holiday.

Table 1. 2006/11 of Calendar (Japan)

Sun	Mon	Tue	Wed	Thu	Fri	Sat
			1	2	*3	4
5	6	7	8	9	10	11
12	13	14	15	16	17	18
19	20	21	22	*23	24	25
26	27	28	29	30		

4.3 Overall Result

The acquired data was registered and analyzed to "Microsoft SQL Server 2000" which is a database management system. From the obtained result, fundamental statistical information is shown below.

4.3.1 All the Acquisition Data Numbers

12,635,568 events

4.3.2 The Number of Electric Pots (= the Number of Users)

3,172 units

4.3.3 Use Days Distribution Classified by Electric Pot

Moreover, it is as follows when it displays in calendar form.

Table 2. Use days - number of electric pots

Use days	#pots	Use days	#pots	Use days	#pots
30	2492	20	13	10	6
29	153	19	9	9	11
28	79	18	17	8	2
27	42	17	11	7	3
26	35	16	4	6	9
25	30	15	9	5	11
24	25	14	8	4	8
23	12	13	4	3	8
22	15	12	6	2	8
21	19	11	8	1	115

Note: Use days 30 = he/she used a pot everyday for 1 month.

Table 3. Number of electric pots(Calendar form)

Sun	Mon	Tue	Wed	Thu	Fri	Sat	
			1	2	*3	4	**Day**
			2905	2905	2904	2885	**#ope**
5	6	7	8	9	10	11	**Day**
2894	2909	2897	2909	2916	2911	2913	**#ope**
12	13	14	15	16	17	18	**Day**
2914	3016	2937	2922	2937	2929	2929	**#ope**
19	20	21	22	*23	24	25	**Day**
2920	2930	2929	2926	2931	2932	2932	**#ope**
26	27	28	29	30			**Day**
2940	2941	2947	2943	2944			**#ope**

Note:
The number of one-day average operation 2,924 sets
The number of the one-day maximum operation 3,016 sets
The number of the one-day minimum operation 2,885 sets

Here, the concrete electric pot use day in 115 electric pots which use only one day in November was as follows.

Table 4. Day - Number of electric pots

Day	11/01	11/13	11/14	11/16	11/20	11/30
#ope	2	100	8	1	1	3

4.4 The Result According to Electric Pot

The statistics result according to electric pot is shown below.

4.4.1 Number of Times of Hot-Water Supply Event by Power Supply Event

The following is the whole number-of-times statistics value of hot-water supply for every power supply injection.

Power supply injection (ps)	75,615 ("hot-water supply" is overwrited)
Hot-water supply (hw)	444,169
hw / ps	5.87

4.4.2 The Power Supply Injection Hot-Water Supply Ratio Classified by Electric Pot

A power supply ON (ps) hot-water supply (hw) ratio is taken out for every electric pot.

The statistical information is shown.

Table 5. hw - ps ratio

hw / ps	# of pot	hw / ps	# of pot
< 1.0	98	10.0 - 20.0	213
1.0 - 2.0	224	20.0 - 30.0	109
2.0 - 3.0	207	30.0 - 40.0	85
3.0 - 4.0	224	40.0 - 50.0	73
4.0 - 5.0	219	50.0 - 60.0	65
5.0 - 6.0	196	70.0 - 80.0	38
6.0 - 7.0	151	80.0 - 90.0	28
7.0 - 8.0	102	90.0 - 100.0	32
8.0 - 9.0	84	> 100.0	292
9.0 - 10.0	71		

One example of less than 1.0 electric pot is shown. It is a pattern on November 1 of a certain electric pot.

Table 6. A pot event of 2006/11/01

Time	event	Time	event	Time	event
1:30	ps	6:40	hw	16:30	ps
1:50	hw	9:00	ps	17:00	hw
4:30	ps	12:50	hw	17:50	ps
4:40	hw	13:00	hw	18:10	hw
6:20	ps	13:40	ps		

4.5 Consideration

A characteristic day can be extracted by seeing the number of operation according to day. Though there is a day which uses an electric pot or is not used by people, if it sees on the whole, it will almost be changeless in the number of operation.

It is over 3,016 sets and 3,000 sets only 11/13. When this cause was explored, there is an electric pot currently used only one day during November, and it became clear that 100 sets of them were 11/13. It is thought that it is based on a sudden factor of some kind.

Next, hot-water supply is observed. All the users' number of times of average hot-water supply was 17,624 times, most number of times of hot-water supply was 18,393 times (11/19, Sun), and the smallest number of times of hot-water supply was 16,727 times (11/1, Wed). It seems that a holiday has generally much number of times of hot-water supply.

The combination of a power supply injection and hot-water supply is considered. By this data, when hot-water supply and a power supply injection are in an observation unit, hot-water supply is adopted.

If it is directions of the usual electric pot, water will be supplied first. Next, a power supply is switched on. And, hot-water supply should continue until hot water is lost.

In this case, it is thought that several times per power supply injection of hot-water supply occurs. So, when the number of times of hot-water supply per number of times of a power supply injection is less than 1.0, it is thought that it is a certain special use pattern.

The example of one day of a certain electric pot is shown as the example of representation (Table 6). From this pattern, there is no concept of keeping warm and it is thought that the power is always turned off after use. In this user, it is thought that it is used not as an electric pot but as an electric kettle. It is recommending the electric kettle in which the proposal of the optimal hot-water supply service has a function a power supply being automatically disconnected after boil, to this user.

Thus, if Plan of the following cycle in PDCA is performed taking advantage of these knowledge, it is possible to draft a high-precision plan.

5 Conclusion and Proposal

This paper described the concept of the "In-Situ Ergonomics" method that uses sensors to extend the direct observation method. It was shown that In-Situ

Ergonomics can solve the problems of cost and investigation period associated with the conventional direct observation method.

In-Situ Ergonomics aims to improve the accuracy of the inputs in ergonomics research. Some results from conventional ergonomics research were not necessarily practical due to the inaccuracy of the information at the input stage.

With conventional ergonomics methods, no particular problems are associated with the analysis and outputs after the inputs are complete. Therefore, applying the In-Situ Ergonomics method should enhance the overall level of ergonomics research.

Acknowledgements. I am thankful to Mr. Shinya of ZOJIRUSHI.

References

1. Nakazawa, et al.: Shinrigaku Manual the Observation Method, Kitaohji-Syobo (1997)
2. Shinya, et al.: In-situ ergonomics: proposal for in-situ ergonomics evaluation method, IEA2003 (Korea) (2003)
3. Shinya, et al.: In-Situ Ergonomics: A proposal of product operation information gathering method for ubiquitous computing environmen, HFES2004 (2004)
4. Shinya, Yamaoka.: Datamining of Universal Access Log by using In-Situ Ergonomics Method, HCI International 2005 (2005)

Distributive Medical Management System

Barbara Sorensen[1], Judith Riess[2], and Eric Allely[3]

[1] Air Force Research Laboratory, Mesa, AZ USA
Barbara.sorensen@mesa.afmc.af.mil
[2] Education & Training Solutions, Vienna, VA USA
education-trainingsolutions@cox.net
[3] Tekamah Corporation, Fairfax, VA USA
Eric@allely.com

Abstract. Medical readiness has become a critical concern world wide. Threats across diverse operation environments, coupled with critical voids in both medical management and medical skills has made the development of new training and management tools and operational assessment methods for first-responders a high priority across global communities. That first response role incorporates a broad array of individuals providing field-based emergency medical services. As recently observed during and after the Asian tsunami and Hurricane Katrina disaster in the United States, critical technological tools and training resources were not available to ensure that the involved first responders had pertinent information available to them to properly do their jobs. The deficiencies must be rectified in order to insure rapid response to natural and man made disasters. Tools that distribute knowledge to aid a range of first responders in evaluating a medical situation, guide them in the uniform collection and reporting of critical information and provide telemedicine clinical reach back to medical experts are essential components of today's medical preparedness. The Distributive Medical Curricula Architecture (DMCA) is designed to provide access to "just-in-time" training and streamline assessment and management of remote medical situations across military and civilian environments nationally and internationally.

Keywords: First responder, medical skills training, training management, remote treatment.

1 Introduction

Medical readiness has become a critical concern world wide. New operational threats and diverse operational environments, coupled with critical voids in medical management, skills, and scarce medical resources have made the development of new training, management tools, and operational assessment methods for first-responders a high-priority in both military and civilian communities [1]. As observed during the Asian Tsunami and Hurricane Katrina in the United States, we still need new technological tools and training resources to ensure that first responders have pertinent information available to them in a timely manner to insure rapid response to natural and man made disasters.

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 979–989, 2007.
© Springer-Verlag Berlin Heidelberg 2007

The United States federal, state and local governments as well as the Uniformed Services are investing untold amounts of time and money in the development of first-responder skill sets. However, this approach provides limited focus on curricula development, evaluation of work performance and sustainment training or skills based training. Prior to 9/11, there were few centralized medical controls governing the development and delivery of first-responder training or response, thus, sharing course materials and other training resources in a distributed environment was not a common occurrence. In addition, there were no operating frameworks by which agencies could jointly plan, train and share information. Existing platforms still find it difficult to share data with other platforms even when the units follow the Shareable Content Object Reference Model (SCORM) standard [2], since there is no standardized way of exchanging various course units and training resources. In addition, integration of such platforms in a distributed environment across military and civilian contexts is even more challenging and complicated.

The demand for tools and methods to strengthen remote medical management capabilities has been accelerated. Critical voids in remote operational health care management prompted the development of tools to aid global first-responders in comprehensively evaluating medical situations through a distributive, deployable, protocol-driven training system. Tools that distribute knowledge to aid a range of first-responders in comprehensively evaluating a medical situation, guide the uniform collection and reporting of critical information and provide a telemedicine clinical reach-back to medical experts are essential components of today's medical preparedness and response plans. The DMCA is designed to expand and strengthen existing capabilities by creating a dynamic, interactive, distributed architecture that takes full advantage of current best practices in information science.

Contrary to military actions which are often preemptive to a threat, the civilian or joint civilian/military response to a threat/disaster is usually in reaction to an event. In these situations first responders must have the necessary information and skills available to them to make rapid informed decisions to save human life [3]. The systems, software tools and training aids that are developed should support the learning objective, the standards, behavior and performance needed.

1.1 Goal of the Distributive Medical Curricula Architecture

The development of the DMCA follows a spiral approach and is continually being updated and upgraded in order to meet the overarching goals. The first goal is to provide an architectural foundation for users to build, publish and share internet based services. The second goal is to develop courses, scenarios, simulations and curriculum that provide skill based state-of-the-art training. In addition, DMCA will provide a platform for SCORM compliant training courses, development and delivery of skills based curriculum that is readily accessible and rapidly reconfigurable to support current and evolving requirements for rapid response for military and civilian emergency responders and medical personnel across all services. The final goal is to provide a collaborative environment that can rapidly update and reconfigure course content.

1.2 Structure and Interfaces

While designers of online learning systems have a wide variety of software tools and learning platforms at their disposal, many different vendor tools and platforms are not able to share learning resources and/or interoperate with each other. Software tools should support the learning objective, the standards, behavior and performance needed. The four most common standards are Aviation Industry Computer-based Training Committee (AICC), SCORM, IMS Global Learning Consortium, and Microsoft LRN. The DMCA supports these standards and has developed learning metadata schemes, course structures and software interfaces that provide interoperability between applications and learning resources. The DMCA enables an institution to deliver skills based training — quickly, effectively, with full tracking and reporting — in an adaptable and straightforward manner. The DMCA has the flexibility to link to any page or document across the web, providing links to external reference materials, glossaries, documents and related websites. Bookmark links allow the user to connect to any slide within a presentation and create non-linear navigation within the content. Bookmarks also provide links to external cataloged library documents.

2 Background

The Distributed Medical Curricula Architecture (DMCA) was developed through the Office of the Secretary of Defense (OSD) Small Business Innovative Research program. During the initial investigative phase, the research uncovered a plethora of first responder training from traditional classroom instruction to online courses. However, finding and gaining access to the courses was a difficult task. The early study confirmed that there were many nonstandard and sub standard patient assessment/treatment protocols in use by first responders and that many courses needed to be revised or updated. During extensive interviews with corpsmen and medics we found discrepancies and disconnects between core training and sustainment training and that first responders had limited accessible resources. In some instances the on-line training was misleading in that it was confined to resource centers and not accessible via personal computers.

2.1 Global Treatment Protocols Via Advanced Distributed Learning

The initial approach for the program was to design and develop a distributive and deployable training module that would teach military and counterpart civilian medical personnel an approved treatment protocol for interacting with physicians and other health care professionals. There were two problem statements that we originally addressed. First, because treatment protocols are not standardized across the US military and civilian community there is an undue hardship caused to both the patient and medical personnel. Upon further investigation of medical documentation, critical incident reports, data collected from health care personnel and a search of pertinent literature, it was determined that problems relating to the lack of standardized training were most likely indicative of larger issues in terms of the organization and approach to the design of emergency responder training programs. Driving towards a single

standard, however, was not appropriate for us to undertake, given the diverse nature of the emergency responder community and their missions especially between military and civilian entities.

The second problem addressed was that medical personnel at all echelons lack formal protocol training for assessing patients and reporting patient conditions to physicians that would allow the latter to make the best possible patient diagnosis and direct the medical personnel to provide the best possible initial treatment until the patient can be transported to a medical facility. Overall, it was assumed that a lack of standards resulted in a reduction in the quality of patient care. However, there is currently a lack of patient outcome data, in both military and civilian communities, that makes the direct correlation between a training course, and a particular treatment, with a patient outcome. There has been some outcomes research but more needs to be done to document and report outcomes as they relate to remote medical treatment and operations. We very quickly learned from our surveys and interviews that while we could develop a training model that would teach certain protocols it would not be useful to the entire community. More importantly, it by no means would create standardization. The problem in the emergency response community is not that there is a lack of standards but a plethora of standards and medical personnel are trained to different protocols. Yet there are few centralized medical controls governing the development and delivery of first-responder training.

Most healthcare teaching/learning platforms, where available, are stand-alone systems. They basically provide only one service (teaching/learning) for a given set of units (courses) which are stored within a single system generally as a PowerPoint, MS Word or PDF file. Sharing course materials and other training resources in a distributed environment is not a common occurrence. Existing platforms find it difficult to share data with other platforms, even when the units follow the SCORM standard, as there is no standardized way of exchanging various course units and training resources. In addition, integration of such platforms in a distributed environment is even more challenging and complicated. In the US there are approximately 15 million full and part-time emergency responders, which include fire personnel, nurses, civilian and military law enforcement personnel, Emergency Medical Technicians (EMTs), Emergency Room Physicians, military corpsmen, medics and millions of ancillary medical personnel. These individuals basically represent the traditional front line of First Responders who must have some type of formal (and coordinated) training. This force, in turn, must integrate with community health workers in schools and offices, with public health facilities, security guards, military medics, National Guard units, and others depending on the catastrophe.

3 Distributed Medical Curricula Architecture (DMCA)

The results of the initial study confirmed the need to develop an infrastructure which facilitates the sharing and real time updating of curricula, best practices, lessons learned and outcomes. Research analysis indicated the need for an architecture that supports the development of adaptable training solutions, acts as a knowledge repository and leverages processes and products that enhance skills-based medical preparedness, as well as provide a collaborative environment to review, revise and

develop curriculum. Therefore, the second phase or spiral was based on developing a Distributed Medical Curricula Architecture (DMCA) that supports first responder readiness through training in a distributed environment and takes advantage of developing technologies to guarantee that the first responder has the information and skills necessary to make rapid informed decisions when treating casualties.

3.1 DMCA Design

The DMCA is designed to address key distance learning issues including: Discovery, Access, Integration, Interoperability and Standardization. This approach supports first responders across jurisdictions and disciplines and improves Communication, Collaboration and Coordination. Specific DMCA goals are to: solve discovery and access issues for first responders; provide a standardized method for the organization and distribution of first responder training; and create an infrastructure for measuring and tracking first responder readiness. It is anticipated that the DMCA will create a "community of practice" for the first responder training community and thereby enhance standardization of first responder training initiatives. The best initiatives will be adopted by other institutions. The DMCA will also provide the means to train to specific skills, measure and track first responder skills levels and provide the venue to develop "just in time" courses to meet urgent training needs. These critical training needs will support a variety of users that include: training solutions providers, content developers, publishers, courseware developers, instructors, students, researchers and medical planners and analysts in both military and civilian communities.

3.2 Architecture

The DMCA architecture is an extensible, agile architecture built on current best practices in software engineering and information management. The DMCA is based on Web Services architecture. The use of Web Services is expanding rapidly as the need for application-to-application communication and interoperability grows. Web Services fundamentally do not represent a new concept but a new approach to an old problem – that of enabling disparate applications to talk to one another. This architecture is already a proven force in e-business and the information technology and tele-communications industries. While other training technologies have been proposed and successfully implemented, albeit sporadically, to improve integration in the training community, they do not address the full scope of the problems facing the first-responder community. The DMCA is not meant to replace these previous efforts but to expand on existing capabilities by creating a dynamic, interactive, distributed platform.

3.3 DMCA Key Features and Capabilities

The DMCA will provide a distributed environment across multiple platforms and organizational settings as shown in Figure 1. Depending on the user's needs and their access privileges, the DMCA will allow retrieving or publishing information or resources in a web service framework. Some of the key features include:

- Access to various medical web services that facilitate medical decisions
- Access to key medical procedures and protocols that support first responders

- Access to DMCA's knowledge repositories
- Access to general medical educational resources
- Provide trusted infrastructure for publishers to publish their treatment protocols, tools, and curricula
- Access to just-in-time medical curricula and critical information
- Provide metadata searching across multiple organizations
- Trusted information search and retrieval capabilities
- Exchange reusable learning and teaching resources
- Access to updated products and clinical decision making tools.

The DMCA provides a framework for collaboration allowing curriculum developers to independently create, catalog and share their training materials with others. The framework will simplify the process of discovery, comparison and evaluation for the users of the materials while not requiring conformity to a single training mode or technology. The DMCA will also act as a knowledge repository and clearinghouse for the data and training curriculum with easy access for all interested parties. Our intent is to provide a platform that will have a significant impact on the first-responder community. Through the Web Services capability, DMCA will provide (a) a publishing model to support the coordinated & collaborative development of medical education and training materials, (b) a consumer model for first-responders to search, discover, and "bind" to medical educational and training materials, and (c) a transaction model to support application-to-application communication (e.g., billing processes, database queries and transactions). It is important to note that the DMCA is not necessarily serving the specific educational content to the student, but rather acting as a broker or intermediary between parties, i.e., in this case, the student and the author. Students come to the DMCA looking for educational content, and the DMCA then connects the student with the desired content. After such a connection is established, the role of the DMCA is largely secondary.

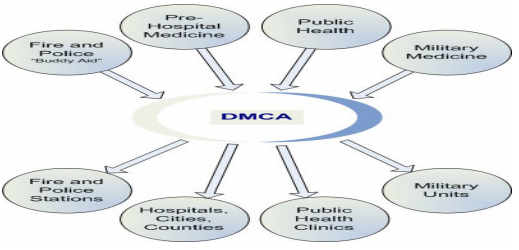


Fig. 1. The DMCA Collaborative Environment

3.4 Demonstrating the DMCA: Integration with Raven Wedge

Based on the technical development and supporting content and organizational research for the DMCA we were prepared to integrate the DMCA with a major first responder exercise in the summer of 2005. The selected exercise, Raven Wedge, was a joint Maryland National Guard (MDNG) exercise scheduled for August 2005. The Maryland Air National Guard was to demonstrate the capabilities of an EMEDS unit for Homeland Security requirements. The Maryland Army National Guard was to demonstrate supporting medical and logistic capabilities necessary for EMEDS units to function in a dirty environment (biological agent). The Public Health Service supported both groups. The Federal Emergency Management Agency (FEMA) [4] Region 3 Raven Wedge exercise was held at Camp Fretterd, Maryland. It was the first

time the 300 participants, primarily from Air National Guard, Army National Guard and Public Health Service had the opportunity to participate in pre and post skills based training geared to the specific exercise.

The scientific and technical development of the DMCA for Raven Wedge followed specific design and planning efforts.

- Member of the team attended planning meetings for the Raven Wedge Exercise from January-August, 2005. At the January meeting we presented the DMCA concept and reviewed initial outline of content with exercise planners, developed draft template for content and established points of contact with subject matter experts. Throughout the next seven months content was added and any problems were discussed at the meetings. In July we opened a site for use by Raven Wedge participants. In addition to "at home" internet access, participants had the opportunity to use the system while on campus during the exercise – the distance learning center at Camp Fretterd was made available. We also completed audio recording of the Chemical and Biological Defense Primer (CBDP). It was implemented as part of the "just in time" training for select Raven Wedge participants. During July we completed final testing of prototype (and Raven Wedge material) using multiple hardware platforms and multiple web browsers.

The DMCA functionality was reviewed with Raven Wedge leadership. Content development for targeted training – especially medical response to a radiological dispersal device (RDD) was completed. Discussions about the variety and formatting of materials to be offered for Raven Wedge (so that we had a good cross section and covered all the necessary learning objectives) were held. We cataloged the materials based on skill categories. Different genre of training materials, as well as varying the type and length of materials to be offered were selected so that the planned scenario was fully covered. Scenario and training materials for Raven Wedge were sent to Laerdal and METI. They participated in the live exercise and played key roles in the field exercise. We played a key role in selecting the courses and instructors for the classroom training to be held the week prior to the three day exercise.

- Selected team members met with senior leadership at the Uniformed Services University of the Health Sciences (USUHS) on the DMCA in support of Raven Wedge demonstration. The administration was fully supportive and authorized subject matter expert (SME) time in support of the effort. The USUHS provided subject matter expertise and collaborative funding for pre-exercise training and post-exercise analysis.

Selected content materials were integrated into a pilot version of the DMCA. These materials included training and evaluation tools for a variety of first responders (EMTs, doctors, nurses, support personnel, etc.). The DMCA was used to distribute pre-exercise training materials to a select group of the first responders participating in Raven Wedge. Pre- and post exercise tests were administered using the DMCA. During Raven Wedge, responder performance was evaluated and after the exercise the DMCA was used to administer post-exercise evaluations and to study the retention of exercise specific training. Raven Wedge was a unique opportunity to test and demonstrate the DMCA and was ideal for our purposes as it was designed to provide a variety of first responders with a wide range of medical skill levels. It was also

fortuitous for many of the participants worked together weeks later in New Orleans on the Katrina relief effort.

4 DMCA Enhancement

The program received continuing developmental funds through an SBIR enhancement effort with matched funding provided by OSD for efforts that integrate the DMCA with other government agencies' programs. Funds were received for a joint project with the Uniformed Services University of Health Sciences (USUHS) [5], Online Preparedness Education Program (OPEP). Both benefit from a jointly funded effort that furthers the development of medical distance learning technologies and OPEP content development for medical personnel and the general public.

4.1 USUHS Online Preparedness Education Program

OPEP is a multidisciplinary, interactive, tiered activity that awards a certificate of completion and continuing education credits. The program provides discipline-specific content for EMT/Paramedics, nurses, physicians, hospital administrators, and emergency managers. The curriculum is designed to assist providers in: recognizing that a WMD event has occurred; managing the medical consequences of an incident; obtaining and utilizing the appropriate resources from internal and external sources; preparing and responding in collaboration with other professions and disciplines; utilizing an incident management system; protecting providers and facilities; and understanding the hierarchical response system. The tiered program consists of a family of activities including: a quick reference a.k.a. "Field Guide" providing rapid access to key management guidelines; a general library of information regarding all aspects of threat agents, cross-cutting issues; and certificated on-line interactive courses. The major tasks for the DMCA enhancement are provided in Table 1.

4.2 DMCA Technical System

DMCA is an online training and tracking system that can utilize almost any kind of educational content, from individual PDF documents to standard web sites to SCORM formatted courses. It can also provide registration and tracking for offline training events and courses. This makes DMCA more flexible than standard LMS systems which typically support content formatted either in SCORM or possibly a proprietary format. The flexibility of DMCA results from its modular design. The main module manages a database of available learning content, including how to access each content type. The database also contains student information and tracks the content they have viewed, and the training results they have achieved. The main module itself does not contain functionality to display the learning content. It only stores information on how the student can be directed to the content, and provides functionality to the other modules to report back the students' progress and scores.

Additional modules manage the student display of SCORM content, displaying and tracking the student's performance in quizzes and exams, and other content available to the students. When students complete the content provided by the additional modules, the modules typically report students' progress and scores back to the main

module. If a remote module (such as a third party web site) does not provide feedback on students' progress and completion status, the content is marked completed the first time the student launches it. Naturally only non-scored content like slideshow presentations or PDF documents are suitable to be accessed in this manner.

Table 1. DMCA Enhancement Tasks

Upgrade of User Testing and Survey Tools	The testing and survey tools upgrade includes the development of an improved web interface for quiz authoring and editing; a downloadable, stand-alone JAVA based quiz editor; support of multiple choice /multiple answer questions; and implementation of question shuffling and random question selection options. The improved testing of the new survey tools will: - Provide enhanced web based support for assessing student performance by providing a platform for the development, delivery and management of online tests and surveys. - Provide expanded question type/function for multiple choice/multiple answer (Multiple-choice questions provide opportunity to offer feedback that gives basis for correct answer). - Implement question shuffling and random question selection permitting the generation of skill based tests for self-administered proficiency testing - Allow instructors, designers and publishers to work off-line.
Upgrade of Learning Management System	A general upgrade of the Learning Management System (LMS) supports both USUHS and the wider OSD community needs in distance learning applications. LMS functionality would be improved to enable additional (more detailed) book marking as well as user "time tracking" to provide information about the length of time users spend on each page.
Enhanced Communication	Enhanced communications include: - Email notifications for the user on pending or expired certificates; new course offerings (filtered by user profile), and "Current Events" postings. Enhanced communications for the administrator/institution would include: - expanded course statistics, user feedback (by course), survey results, and certification/documentation requests.
Creation of 'Current Events' Portal	The integration of a "Current Events" portal as part of the USUSH WMD Distance Learning homepage includes the ability to publish information from other related sites and permit filtering of information by user profile. The "current events" section would include a link checking tool to identify "broken" or "dead" links on the website. "Current Events" links for external and internal events will be filtered based on profiles of the individual user. This task will include the creation of a publishing interface for current events.
Upgrade of Administration Interface	The administrative interface is being enhanced for more general use and will include tools for developing and deploying user survey tools and item analysis features for question responses (applicable to user tests and user surveys).
Content Development Support	Content development is guided by CDHAM (USUHS) and will focus on Chemical Agents. Funding will support the development of Field Guide(s), Library References and an On-line Course addressing Chemical Agents as weapons of mass destruction and will parallel the current materials focusing on radiological weapons.
System Support and Maintenance	System Support and Maintenance includes 24/7 Internet user access, network hardware maintenance, system backup and archiving, server administration, minor software updates and "bug fixes" and direct support of USUHS WMD Distance Learning administrators and content developers

Currently DMCA utilizes two special purpose standalone modules that directly interface with the main module (Fig 2). One is a SCORM compliant LMS (Learning Management System) that displays any SCORM compatible content to the students. The other is a quiz server that displays and tracks the students' performance in quizzes and exams the standalone modules provide simple web service interface to allow the main module to register existing students in the module database.

The web interface also provides a method for the main module to pass authenticated students to specific content in the other modules. The standalone modules provide simple web service interface to allow the main module to register

existing students in the module database. The web interface also provides a method for the main module to pass authenticated students to specific content in the other modules without requiring an additional authentication step. The main module in turn provides a method for the other modules to report content completion status and scores back to the main tracking database. For tracking completion and scores of offline training, the main module contains forms for the teachers to enter student progress and scores. Currently, DMCA supports student, registrar, and system administrator roles. Registrar users can view the students' training results and progress. System administrators are responsible for maintaining the global configuration data for the system. In the future the role of teacher/publisher is planned to be added. Teacher/publishers would be allowed to publish online or offline training content to the DMCA web site, and in case of offline content, they would be allowed to enter student scores into the system.

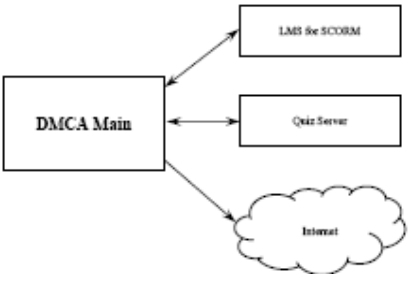


Fig. 2. DMCA format

5 Conclusion

There are a broad range of stakeholders and first-responder groups across military and civilian organizations that share a common problem – lack of medical readiness. The DMCA is one part of the solution that would provide a technical architecture with end-to-end solutions for training out emergency responders as well as providing an intelligent knowledge repository. An easily accessible repository could be leveraged and mined across multidisciplinary domains by medical planners, analysts and researchers. The DMCA could also facilitate the development, sharing, and real-time updating of curricula, best practices, lessons learned and outcomes.

Although building the DMCA presents challenges, the utility and capability to assist in educating and training our emergency response personnel is well worth the endeavor. Standards for procedures, skills development, performance measures, interagency interaction and funding are the biggest challenges. DMCA is currently working on these issues as an integral part of future development.

References

1. Future of Emergency Care in the US Health System, Institute of Medicine of the National Sciences, National Academic Press, Wash D.C. (June 2006)
2. Sharable Content Object Reference Model (SCORM 2004) 3rd. edn. The ADL Technical Team. Sponsored by the Office of the Secretary of Defense for Personnel and Readiness (November 2006)

3. Knudsen, D., Johnson, J.A., Ledlow, G.R., Barbisch, D., Binns, E.: Are you ready to Respond? EMS.com Responder (2007)
4. Emergency Response Team-Mitigations Operations Manual, Federal Emergency Management Agency (March 2000)
5. Wayne, B.: Department of military and emergency medicine: Military studies I, Combat medical skills course handbook. Bethesda, MD: Uniform Services University of the Health Sciences (2003)

Basic Consideration for Various Interfaces from Viewpoint of Service

A. Toshiki Yamaoka and Mayuko Yoshida

Wakayama University, Faculty of Systems Engineering,
Sakedani-930, Wakayama city, Japan
yamaoka@sys.wakayama-u.ac.jp

Abstract. Service business was classified into three categories. A design framework of user interface, especially screen design for the three categories of service was constructed roughly from viewpoint of theory of regulatory focus, transactional analysis (TA) and psychological impression. According to the three categories, three examples of screen design were designed roughly.

Keywords: Service, screen design, user interface, theory of regulatory focus, transactional analysis (TA), psychological impression.

1 Introduction

Service business in 21st century becomes very important industry. However, relationship between service and user interface (screen design) has not been examined until now. Then, the paper describes a design framework of user interface, especially screen design for service from viewpoint of theory of regulatory focus, transactional analysis (TA) and psychological impression.

2 Classification of Service

61 business persons were asked regarding objects and services. They are as follows.
Objects:

1. medical treatment
2. restaurant
3. self-restaurant
4. laundry
5. laundry(self service)
6. beauty salon
7. movie/drama
8. ATM
9. plane
10. reservation of plane ticket
11. home delivery service

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 990–996, 2007.
© Springer-Verlag Berlin Heidelberg 2007

12.hotel
13.amusement park
14.game arcade
15.education
16.repair of home electrical appliance
17.net shopping
18.retrieval in bookstore

Services:

1. attitude of employee
2. good HMI (Human Machine Interface)
3. changed service based on customer
4. providing reliable service
5. providing effective service
6. expected service accepted
7. providing added service
8. staff cooperated with customer for service
9. low price
10.taking much time for service
11.enjoyable process of service

If object and service are related, subject checked. The data gotten from questionnaire were analyzed using correspondence analysis and cluster analysis.

Table 1. Relationship between objects and services (frequency / total)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
medical treatment	0.70	0.23	0.60	0.91	0.55	0.75	0.23	0.26	0.32	0.83	0.04
restaurant	0.87	0.06	0.36	0.40	0.59	0.77	0.51	0.08	0.25	0.08	0.42
self-restaurant	0.15	0.51	0.06	0.30	0.64	0.34	0.09	0.26	0.83	0.02	0.09
laundry	0.38	0.05	0.12	0.57	0.52	0.67	0.19	0.14	0.45	0.19	0.00
laundry (self service)	0.00	0.62	0.00	0.35	0.52	0.35	0.07	0.31	0.76	0.03	0.03
beauty salon	0.09	0.02	0.68	0.71	0.32	0.78	0.51	0.12	0.22	0.59	0.44
movie/drama	0.20	0.11	0.04	0.22	0.30	0.44	0.20	0.07	0.17	0.00	0.44
ATM	0.20	0.75	0.10	0.72	0.73	0.49	0.04	0.31	0.00	0.18	0.00
Plane	0.76	0.22	0.37	0.56	0.37	0.53	0.45	0.08	0.14	0.06	0.22
reservation of plane ticket	0.19	0.61	0.13	0.61	0.74	0.45	0.03	0.39	0.00	0.03	0.00
home delivery service	0.52	0.07	0.17	0.71	0.57	0.57	0.11	0.15	0.44	0.07	0.00
hotel	0.97	0.10	0.53	0.65	0.41	0.73	0.80	0.06	0.16	0.12	0.43
amusement park	0.44	0.33	0.19	0.26	0.43	0.43	0.41	0.26	0.19	0.05	0.64
game arcade	0.13	0.44	0.08	0.22	0.25	0.28	0.19	0.31	0.42	0.03	0.44
education	0.52	0.05	0.55	0.67	0.26	0.60	0.26	0.19	0.19	0.57	0.36
repair of home electrical appliance	0.61	0.06	0.26	0.68	0.52	0.61	0.13	0.06	0.39	0.26	0.00
net shopping	0.11	0.63	0.13	0.66	0.66	0.47	0.34	0.24	0.63	0.05	0.21
retrieval in bookstore	0.10	0.67	0.08	0.23	0.69	0.49	0.05	0.39	0.00	0.00	0.10

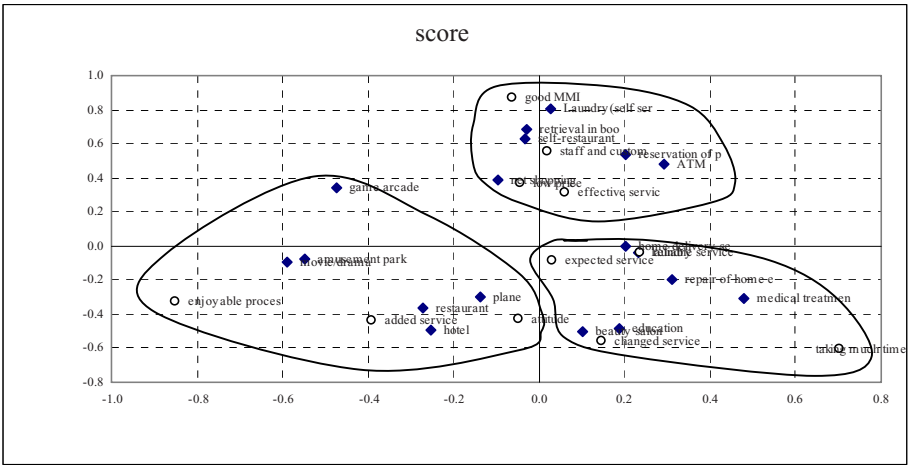


Fig. 1. Three classification of service

The questionnaire was analyzed using correspondence analysis and cluster analysis. The results are as follows.

1. "pleasure" service group
   Objects: restaurant, movie/drama, plane, hotel, amusement park, game arcade
   Services: providing added service, enjoyable process of service

2. "low cost and effective" service group
   Objects: self-restaurant, laundry (self service), ATM, reservation of plane ticket, net shopping, retrieval in bookstore
   Services: good HMI (Human Machine Interface), providing effective service, staff cooperated with customer for service, low price

3. "reliable and expected" service group
   Objects: medical treatment, laundry, beauty salon, home delivery service, education, repair of home electrical appliance
   Services: attitude of employee, changed service based on customer, providing reliable service, expected service accepted, taking much time for service

3 Characteristics of Three Classification of Service

The three classification of service was analyzed using transactional analysis (TA) theory of regulatory focus and psychological impression after using.

Theory of regulatory focus consists of promotion focus and prevention focus [1] [4]. The promotion focus means that every person works to maximize positive action results and minimize non-positive loss for promotion and self-realization. The prevention focus means that every person works to protect themselves against the

occurrence of unpleasant states. So, user interface of product and service and so on can be designed and analyzed based on the viewpoint of promotion focus and prevention focus.

1. “pleasure” service group : promotion focus
2. “low cost and effective” service group: prevention focus
3. “reliable & expected” service group: prevention focus

Table 2. characteristics of three classification of service

group	(1) “pleasure” service group	(2) “low cost and effective” service group	(3) “reliable & expected” service group
object / user interface	restaurant, movie/drama , plane, hotel, amusement park, game arcade	self-restaurant , laundry(self service), ATM, reservation of plane ticket, net shopping, retrieval in bookstore	medical treatment, laundry, beauty salon, home delivery service, education, repair of home electrical appliance
	user interface	user interface	user interface
theory of regulatory focus	promotion focus	prevention focus	prevention focus
transactional analysis (**TA)**	A →A(adult) C → C(child) P(parent) → C	A →A	A → A
psychological impression after using	feel pleasure or impression	feel ease to use	feel sense of security or reliability

As TA [3] is a method to improve communication, relationship between human and machine (system) was analyzed using TA.

1. “pleasure” service group:
   Device provides human with information based on three different communication patters of TA which is “adult to adult”, “child to child” and “parent to child”.

2. “low cost and effective” service group:
   Device provides human with information based on “adult to adult” pattern of TA.

3. “reliable & expected” service group
   Device provides human with information based on “adult to adult” pattern of TA.

Finally, the three classification of service are analyzed from viewpoint of psychological impression [2] after using.

1. “pleasure” service group : feel pleasure or impression
2. “low cost and effective” service group: feel ease to use
3. “reliable & expected” service group: feel sense of security or reliability

The three classification of service were made clear using transactional analysis (TA) theory of regulatory focus and psychological impression after using.

4 A Screen Design Example Corresponded to Three Classification of Service

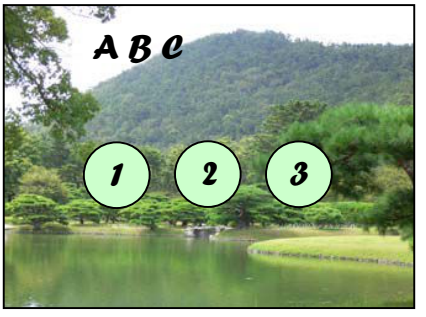
Specifications of screen design based on characteristics of three classification of service are shown in table3.

Table 3. a specification of screen design based on characteristics of three classification of service

group	(1) “pleasure” service group	(2) “low cost and effective” service group	(3) “reliable & expected” service group
	user interface	user interface	user interface
theory of regulatory focus	promotion focus	prevention focus	prevention focus
transactional Analysis (**TA**)	A →A(adult) C → C(child) P(parent) → C	A →A	A → A
psychological impression after using	feel pleasure or impression	feel ease to use	feel sense of security or reliability
title	needful	needful	needful
explanation	describe decoratively	describe simply	describe in detail
parts (button, etc)	decorative, no shade of button	simple, need shade of button	metaphor
confirmation of operation	needless	needless (automation) (sometimes, needful)	needful
concrete / abstract	concrete	abstract	concrete

The following design ideas are created according to table 3.

1. “pleasure” service group :



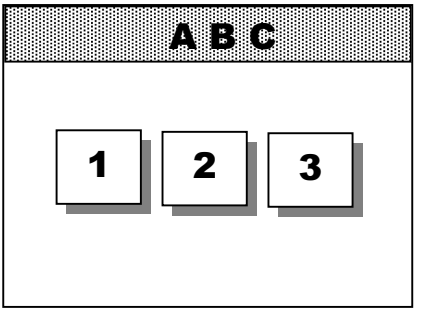
Theory of regulatory focus:
promotion focus

Transactional Analysis (**TA**):
C C

Psychological impression:
feel pleasure

Design: concrete

2. “low cost and effective” service group:



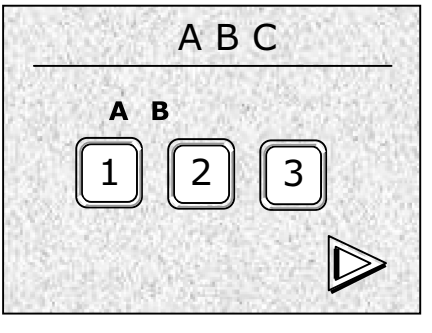
Theory of regulatory focus:
prevention focus

Transactional Analysis (**TA**):
A A

Psychological impression:
feel ease to use

Design: abstract
The shade of button is needed.

3. “reliable & expected” service group



Theory of regulatory focus:
prevention focus

Transactional Analysis (**TA**):
A A

Psychological impression:
feel sense of security

Design: concrete
The button is quoted from real button.
Confirmation of operation is needed.

5 Discussion

Framework of user interface, especially screen design based on service was made clear roughly. Although service was classified into three categories, the concept of

service has to be identified. The layout, parts and so on of screen design was decided according to the identified concept. The design procedure is as follows.

1. Examine theory of regulatory focus.
   Promotion focus or prevention focus is decided according to the theory of regulatory focus.

2. Decide type of communication between human and machine (system).
   The combination of adult, child and parent is decided based on the promotion focus or prevention focus.

3. Psychological impression is decided.
   The psychological impression after using a machine is decided according to TA.

References

1. Higgins, E.T.: Beyond pleasure and pain. American Psychologist 52, 1280–1300 (1997)
2. Kobayashi, Y., Katho, Y., Nishida, Y., Mizuno, Y., Yoshida, K., Yoshida, M., Yamaoka, T.: Consideration of product evaluation method focused on interaction between product and user. In: Proceedings of Knsai Chapter of Japan Ergonomics Society, pp. 5–8 (2006)
3. Okano, Y., Tada, T.: Start for new oneself – TA for management, syakai-sangyou-kyouiku-kenkyuujo (197)
4. Werth, L., Foerster, J.: How regulatory focus influence consumer behavior. European Journal of Social Psychology 36, 1–19 (2006)

Specialized Design of Web Search Engine for the Blind People

Yi-Fan Yang[1] and Sheue-Ling Hwang[2]

[1] HTC, Taipei research center, Hsin-Tein, Taipei, Taiwan
[2] Department of Industrial Engineering and Engineering Management, National Tsing Hua University, Hsinchu, Taiwan
slhwang@ie.nthu.edu.tw

Abstract. In this research, Google Website was taken as a Web prototype to get modified. Specialized Search Engine for the Blind (SSEB) was constructed with an accessible interface and some improved functions (i.e., searching assistance functions, user-centered functions, and specialized design for the blind). An experiment was conducted to verify the effect of SSEB. From experimental results, it's significant that users attained better performance in SSEB than in Google. Users also showed higher satisfaction with SSEB.This research proves that the consideration for designing an accessible website for the blind users is very important. It is expected that the users group of SSEB can be expanded to all visual-impaired people in the future. So that all people can keep pace with the changing World Wide Web, and make good use of all the Internet resources without disabilities and limitations.

Keywords: Web search engine, Blind; Web searching, Web accessibility, Google.

1 Introduction

Accessibility is becoming increasingly indispensable to the Internet experience. For the disabled persons, especially the blind people, the minimum condition is to ensure that everyone can understand all the content of the webpage.

World Wide Web Consortium [11] developed "Web Content Accessibility Guidelines 1.0" which contains 14 significant guidelines. These guidelines explain how to make Web content accessible to people with disabilities. They are intended for all Web content developers and for developers of authoring tools.

In compliance with the difficulties encountered by the Department of Communication and the Arts, Online Public Access Initiative, and some studies of the people with disabilities in surfing the web [3][10], we can classify the difficulties according to the computer equipment and the data obtaining operation.

Several studies about audio supplementary instruments have been great contributions to the people with visual impairment [2][4][7].

The Web Access Project [4] has developed methods for adding captions and audio descriptions to movie clips, making Web-based multimedia more accessible to users with sensory impairments.

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 997–1005, 2007.
© Springer-Verlag Berlin Heidelberg 2007

World Wide Web Consortium [12] recommended that content developers should offer search mechanisms which satisfy varying skill levels and preferences when providing search functionality. Most search facilities require the user to enter keywords for search terms. Users with spelling disabilities and users unfamiliar with the language of the web site will have a difficult time finding what they need if the search requires perfect spelling. Search engines might include a spell checker, offer "best guess" alternatives, query-by-example searches, similarity searches, etc.

The Web and especially major Web search engines are essential tools in the quest to locate online information for many people. Jansen & Spink [9] examined characteristics and changes in Web searching from nine studies of five Web search engines based in the US and Europe. They compared interactions occurring between users and Web search engines from the perspectives of session length, query length, query complexity, and content viewed among the Web search engines.

Jansen et al. [8] have analyzed more than 50,000 queries in Excite's query log and found that users use few terms (2.21 per query) when searching the database. The survey also shows that only approximately 5% of the users use advanced search features like the Boolean AND-operator (very few use OR and AND NOT) and relevance feedback (the latter is used in 5% of all queries). A third important point is the examination of search results. Only 20% of the users looked beyond the first two result pages. On average each user looked at 2.35 pages.

Google [6] makes a beta function open to public use, that is, Personalized Search. Using Personalized Search, users can get the results most relevant to them, based on what they've searched for in the past.

In this accessible Web, the functions are constructed from the blind people's requirements. Four main objectives of the research are as below:

(1) To design an accessible Interface especially for the blind people;
(2) To modify the searching process, and to simplify the input methods and steps;
(3) To construct user-centered functions and specialized design for the blind;
(4) To help the blind people to shorten their searching time, and to seek the right information successfully.

It is expected that the blind person can learn and use this website by him/herself just like ordinary person, and get all the information he or she needs by this system. So that they can keep pace with the changing World Wide Web, and improve their learning scope and efficiency.

2 Methods

2.1 Basic Design Structure of SSEB

Google Website[5] is selected as a Web prototype to construct an improved web searching engine: Specialized Search Engine for the Blind (SSEB). To refer to Chapanis'[1] methods of system design, operational-need for the blind people and system requirements are analyzed firstly. Then functions and interface are designed, and the website simulator, i.e. SSEB is constructed. An experiment is conducted to

test and verify the performance of SSEB and Google, and the performance between ordinary and blind people.

In addition to improve the blind users' searching process, some user-centered functions will be added to SSEB to make it more accessible. The interface will be designed to conform to web accessibility especially for the blind people.

Three primary aspects of improvement are as below:

(1) To design an accessible interface especially for the blind people;
(2) To modify searching process;
(3) To construct user-centered functions.

The output of the system is read by audio supplementary instruments mainly. Information on the interface of SSEB will be read by screen readers such as JAWS and Big Eyes. Users use a keyboard to input keywords, keyboard shortcuts, personal setting, or other data. Three major parts of the design functions are:

(1) Analysis of searching conditions;
(2) Analysis of searching results;
(3) Specialized design for the blind.

2.2 Constructing SSEB

After confirming all functions, SSEB is constructed by a web development tool, Macromedia Dreamweaver MX 2004. The documents are written in the type of Asp VBScript, and Microsoft Access is adopted as the database to store the query results and subjects' searching records. A friendly interface is designed by considering the blind people's experience of using Google

Basic Functions

1) Users have to register a user ID in SSEB in order to use other specialized functions.
2) Uses have to login SSEB by their registered ID and password.
3) Users can input keywords to execute the searching function.

New Establishing Functions

1) Adjustable display function

Users can change the display mode according to their preference. Fig. 1 shows the unabridged mode, and each search result contains five items in this mode, i.e. website title, concise content, url, webpage size, and the latest updated time.

In the brief mode, three items of each search result are abridged, and only two are left: website title, concise content. This mode can help users to focus on the focal point of each result, and then users can browse the result pages quickly.

2) Results sorting function

Searching results can be sorting by five items (i.e. website title, concise content, url, webpage size, and the latest updated time). Users have to choose one item, and to ascend or descend the data.

3) Bookmark function

Users can label bookmarks for the sake of retrieving searched websites. If they are interested in any of the results, they can click the link of the serial number above each result, and then this result will be stored into the "Bookmark List" page. With this function, users can browse more searching result pages and find more useful websites for them.

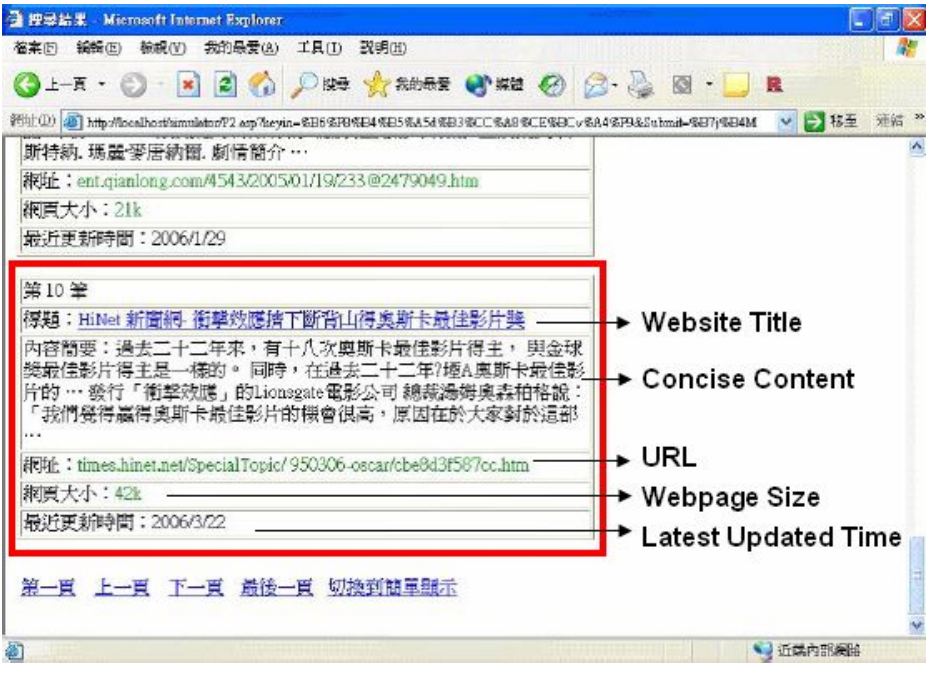


Fig. 1. The unabridged mode

Experiment and Verification. After SSEB had been constructed and accomplished, an experiment was conducted to test the performance of this system.

The experiment was designed to compare the performance of SSEB with the original Google search engine (i.e. control). All subjects were asked to fulfill two kinds of experimental tasks:

(1) Subjective judgment and opinion:

Questionnaires were carried on to test learning efficiency and users' degree of familiarity and preference with this simulator;

(2) Objective searching experience:

Searching time and correct rate were collected after each searching tasks.

Twelve subjects were invited to attend this experiment. Eight of them were visual-impaired people (including four blind people), and four were ordinary people. The experiment was designed so that each of the twelve subjects used the two systems in a

random order. Each subject could complete the evaluations at an individual pace, to ensure that the experiment reflected real-world conditions and to diminish the effects of fatigue on the users.

SSEB was a helper application which ran within Internet Explore of Windows XP on a notebook computer. All the text would be read out by the screen reader (i.e., Big Eyes I). The ordinary subjects were requested to wear an eyeshade during the experiment to imitate the blind people.

Searching time and correct rate would be collected and analyzed to compare the performance of SSEB with Google. The performance of three groups (i.e., the blind people, intermediate visual-impaired people, and the ordinary people) would also be discussed. The results of the subjective questionnaires would be analyzed to modify the system.

3 Results and Discussion

The data were collected from the experiment to compare the performance of SSEB with Google, and to find out if there is any difference of searching achievements between the visual-impaired people and the ordinary people. Subjects' performances were evaluated by three indexes: the searching time, the correct rate of searching tasks, and statistics of the subjective questionnaires.

3.1 The Searching Time

A two-way analysis of variance was adopted to test the equality of populations means classified by two factors: system and group. The two-way ANOVA is a factorial method to study the interaction among the independent variables of system and group and their effect on average searching time. There was a significant difference among the systems (p-value = 0.024 < 0.05). However, the effect of the three groups on the searching time was not significant (p-value = 0.199 > 0.05), and the results of a two-way ANOVA showed no significant interaction between Systems and Groups on average searching time (p-value = 0.163 > 0.05).

Because each participant fulfilled both the searching tasks of the two systems, the testing method of paired t-test was applied to test the mean difference between SSEB and Google. The results indicated that the mean time for SSEB (mean = 393.92) was statistically (p-value = 0.024 < 0.05) lower than the mean time for Google (mean = 717.33). It meant that users spent shorter time on searching the designated answer in SSEB than in the original Google search engine.

In order to find out the difference in detail, the searching time of three groups were analyzed separately. One-way ANOVA was adopted to test the equality of population means for SSEB and Google of each group. The results indicated that the mean time for SSEB of Group3 (Ordinary; mean = 403.50) was statistically (p-value = 0.027 < 0.05) lower than the mean time for Google of Group3 (mean = 1050.75). In other words, the ordinary people had better searching performance in SSEB.

There was no significant difference (p-value = 0.269 > 0.05) between the searching time in SSEB (mean = 280.25) and Google (mean = 601.00) of Group1 (Blind), and

the effect of the two systems on Group2 (Intermediate) was not significant (p-value = 0.991 > 0.05) either (SSEB mean = 498.00 and Google mean = 500.25).

3.2 The Correct Rate

It was found that the data of correct rate were not distributed in a specific distribution, so Friedman test, a nonparametric analysis of a randomized block experiment, was adopted to compare two treatments: SSEB and Google. Subjects were taken as blocks, and there were twelve duplications. (Suppose that there was no interaction between treatment and block.) All treatment effects were not rejected as zero (p-value = 0.414 > 0.05); that is, the effect of the correct rate of SSEB (mean = 0.8667) and Google (mean = 0.7250) was not significant.

To analyze the effects of different groups, Friedman test couldn't be adopted because there was no block. Kruskal-Wallis (Distribution-Free Test) was taken to compare three treatments: Blind (Group1), Intermediate (Group2), and Ordinary (Group3). An assumption for this test was that the samples from the different populations were independent random samples from continuous distributions, with the distributions having the same shape.

Three population medians were not significantly different (p-value = 0.939 > 0.05). In other words, the effect of the correct rate of three groups was not significant.

In order to find out the difference in detail, the correct rate of three groups were analyzed separately. Kruskal-Wallis Test was adopted to test the equality of medians for SSEB and Google of each group.

The results indicated that the correct rate of SSEB (mean = 1.0000) was statistically (p-value = 0.046 < 0.05) higher than the correct rate of Google (mean = 0.5500) in Group1 (Blind). That is to say, the blind people's searching performance of SSEB was better than that of Google.

However, there was no significant difference (p-value = 0.739, 0.617 > 0.05) between the correct rate of SSEB and Google in Group2 and Group3.

3.3 The Analysis of the Subjective Questionnaires

The subjective questionnaires included two aspects: (1) users' preference for the SSEB or Google; (2) users' integral evaluation of SSEB. To integrate all questionnaires, the one-sample Wilcoxon signed rank test was adopted to analyze data.

Subjects agreed that the configuration (i.e., the place of characters, links, and buttons) of SSEB was more accurate than which of Google (p-value = 0.004 < 0.05). It's less possible to lose their orientation when surfing in SSEB than in Google (p-value = 0.002 < 0.05). SSEB's function of bookmark was more suitable for the blind than Google's (p-value = 0.006 < 0.05). The reaction time and stability of SSEB was better than Google (p-value = 0.011 < 0.05). For the novices, to learn the searching operation of SSEB was easier than that of Google (p-value = 0.001 < 0.05). Users could search out the information they wanted more easily in SSEB than in Google (p-value = 0.007 < 0.05).

Subjects agreed that the layout of SSEB was facile (p-value = 0.002 < 0.05); the function of adjustable display was great assistance for the blind people (p-value = 0.026 < 0.05); the function of results sorting was great assistance for the blind people

(p-value = 0.003 < 0.05); the function of bookmark was great assistance for the blind people (p-value = 0.033 < 0.05); the individual specialized functions (e.g., personal calendar, schedule reminding, newsflash, etc.) would be great assistance for the blind people (p-value = 0.002 < 0.05). However, subjects did not show significant favor for the number of results per page (p-value = 0.221 > 0.05).

3.4 Discussion

From the results of the experiment, the improved design of SSEB had shortened all users' searching time obviously. It could also help the blind people to search for the correct object, and enhance the integral performance significantly. Based on Web Content Accessibility Guidelines [11], the main factors that affected the performance were analyzed and described as follows:

1) SSEB provided equivalent alternatives to visual content, i.e. the alternative words of images, graphs, or buttons. Then the assistive technology, e.g., screen reader, would read those alternative words to visual-impaired people, so they could easily understand the information on the web page. Some unnecessary elements in SSEB were even removed in order not to confuse the users. However, Google didn't achieve this guideline, and therefore the difference between the two systems existed.
2) The characters in the pages of search results in Google were not read smoothly by Big Eyes I. By contrast, the characters of SSEB were pronounced unhindered, because the paragraphs were designed well to ensure that pages featuring new technologies transform gracefully.
3) There were many repeated functional items in the forefront of the result pages of Google, which would waste users' time on skipping among those items. The needless items would lower users' searching efficiency. As regard to SSEB, it provided clearer navigation mechanisms. It included document titles and identified the target of each link explicitly.
4) As to interim solutions, SSEB provided default, place-holding characters in edit boxes and text areas, and didn't cause pop-ups or other windows to appear without informing the user.

From the subjective questionnaires, subjects showed apparent preference for SSEB. They agreed that the configuration, navigation concepts, the function of bookmark, reaction time and stability of SSEB were all better than which of Google. Furthermore, it was easier for the novices to learn the searching operation of SSEB than that of Google, and users could search out the information more easily in SSEB than in Google.

As to subjects' integral evaluation of SSEB, their satisfaction with SSEB was quite high. Subjects agreed that the layout of SSEB was facile. The individual specialized functions, the functions of adjustable display, results sorting and bookmark would be a great help for the blind people. The advantages of SSEB were that the structure of pages was well organized and clearly stated, and to learn and operate SSEB was effortless. SSEB could help visual-impaired people to find information more easily and efficiently.

4 Conclusions

On the basis of Web Content Accessibility Guidelines 1.0 (World Wide Web Consortium, 1999), SSEB was constructed from the Google Website as a web prototype. SSEB provided an accessible interface designed especially for the blind people, and users could browse through this system by screen readers. With the friendly system, the blind people could use simple logical thinking to achieve their searching tasks.

With the assistive functions of SSEB (e.g., adjustable display and results sorting), the advanced searching process and conditions were modified, and the searching methods and steps to achieve the same goal were also simplified. The experimental results demonstrate that these functions could help the blind people to shorten their searching time, and to seek the right information successfully.

Especially for a large number of information retrieving, the user-centered functions and specialized design for the blind (e.g., bookmarks, personal calendar, schedule reminding, newsflash, etc.) would be great assistance for the blind people. It would be extremely convenient to the users that they could set SSEB for their personal behavior and habit to satisfy their own demand.

Acknowledgments. The support of the National Science Council, Taiwan (grant NSC 95-2221-E-007-114) is gratefully acknowledged.

References

1. Chapanis, A.: Human Factors in Systems Engineering. John Wiley & Sons, Inc, New York (1996)
2. Chiang, M.F., Cole, R.G., Gupta, S., Kaiser, G.E., Starren, J.B.: Computer and World Wide Web Accessibility by Visually Disabled Patients: Problems and Solutions. Survey of Ophthalmology vol. 50(4) (July-August 2005)
3. Evett, L., Brown, D.: ext formats and web design for visually impaired and dyslexic readers-Clear Text for All. Interacting with Computers 17, 453–472 (2005)
4. Goldberg, L., Freed, G.: Making multimedia accessible on the World Wide Web. Technology and Disability 8, 127–132 (1998)
5. Google (December 15, 2005) (2005) Available http://www.google.com/
6. Google (April 22, 2006) (2006) Personalizing your search results. Available: http://www.google.com/support/bin/topic.py?topic=1593&hlrm=zh_TW
7. James, F.: AHA: audio HTML access. Computer Networks and ISDN Systems 29, 1395–1404 (1997)
8. Jansen, B.J., Spink, A., Saracevic, T.: Real life, real users, and real needs: a study and analysis of user queries on the web. Information Processing & Management 36(2), 207–227 (2000)
9. Jansen, B.J., Spink, A.: How are we searching the World Wide Web? A comparison of nine search engine transaction logs. Information Processing and Management 42, 248–263 (2006)

10. Seiler, R.J., Seiler, A.M., Ireland, J.M., Guy, A.M., Woodward, K.N.: Enhancing Internet access for people with disabilities. Computer Networks and ISDN Systems 30, 739–741 (1998)
11. World Wide Web Consortium Web Content Accessibility Guidelines 1.0. (September 14 2005) (1999) Available http://www.w3.org/TR/WAI-WEBCONTENT
12. World Wide Web Consortium. Core Techniques for Web Content Accessibility Guidelines 1.0. (November 2, 2005) (2000) Available http://www.w3.org/TR/WCAG10-CORE-TECHS/

The Augmented Patient Chart: Seamless Integration of Physical and Digital Artifacts for Hospital Work

Myrna S. Zamarripa[1], Victor M. Gonzalez[2], and Jesus Favela[1]

[1] Department of Computer Science, CICESE Research Center, Mexico
[2] School of Informatics, University of Manchester, United Kingdom
mzamarri@cicese.mx,vmgonz@manchester.ac.uk,favela@cicese.mx

Abstract. Even with the introduction of computer technology, paper-based artifacts remain ubiquitous in hospital settings. The need to manually transfer and update information from the physical to the digital realm is a common practice among hospital staff, which, although usually well managed, at times becomes a source for errors and inconsistencies. This paper presents an augmented patient chart system that preserves the use of paper and allows capturing information directly to the system through the use of a digital pen. An evaluation of the system with 22 volunteers indicates a significant reduction in the number of errors while reading information, a significant increment on the accuracy while annotating data, as well as a trend towards less time spent while annotating data on the digital paper. Based on our results, we argue that the design preserves the advantages associated with paper while increasing the availability of information and its trustworthiness.

Keywords: Ubiquitous computing, Hospital work, Physical artifacts, Digital paper.

1 Introduction

Even with the introduction of computer technology, paper-based artifacts remain ubiquitous in hospital settings. One of the reasons for paper being so prevalent is the high level of mobility experienced by hospital workers, whom have to move from place to place, carrying information on portable artifacts such as notebooks, clipboards, and information sheets [1]. In clear contrast to a desktop-based working environment [2] or central control rooms [5], the information in hospitals is not generally concentrated in a single place but it is distributed in different sites. The storage of information in physical paper artifacts, results on clear inefficiencies, as information cannot be retrieved promptly whenever and by whoever might need it. Furthermore, the need to manually transfer and update information from the physical to the digital realm is becoming a daily practice among hospital staff, which, although usually well managed, at times becomes a source for errors and inconsistencies. Estimates indicate that up to 98,000 people die in the US as a result of medical errors [4], which serves to realize the importance of proper information management in hospital settings.

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 1006–1015, 2007.
© Springer-Verlag Berlin Heidelberg 2007

The use of portable computers and personal digital assistants (PDAs) could help support information management in such a mobile working environment, yet, it is clear that much is lost when paper affordances are left out. For instance, a study comparing the use of paper to PDAs or TabletPCs for the elaboration of medical notes showed that writing is faster when using pen and paper [7]. Consequently, rather than replacing paper, our aim is to use pervasive computing technology to seamlessly integrate paper into physical-digital hospital information systems. This integration of physical artifacts and digital information is one of the main challenges facing the field of ubiquitous computing.

We developed an augmented patient chart system, based on Anoto™ technology that preserves the use of paper and allows capturing information directly to the digital system through the use of a digital pen. The solution maintains the essence of the original patient chart and introduces features that make it richer by providing up to date information and more powerful graphical presentations. The system was designed to facilitate its integration to the information systems operating in a typical hospital (e.g. Electronic Patient Record, Clinical Decision Support Systems, Computerized Physician Order Entry Systems, etc).

The rest of the paper is organized in the following way: Section 1 presents some observations and requirements derived from a case study conducted to inform the design of the application. Section 3 discusses the application, its architecture and functionality. Section 4 presents results from an evaluation of the system by 22 nurse students. Finally, Section 5 closes the paper with our conclusions and discussion of future work.

2 Hospital Patient Charts: A Case Study

Our design emerged from an ethnographic study consisting of interviews and shadowing of people at a mid-size public Hospital in Ensenada, Mexico. We focused our study on operative nurses as they interact constantly with the patient chart. Using the patient chart, the operative nurses are in charge of following up and carrying out the indications left by physicians treating the patients. During three days, two nurses were followed (shadowed) during their entire shift to understand the relevance of the patient chart on executing their daily activities.

As a result of the observations we were able to identify:

- The work processes around the patient chart
- The general structure of the patient chart as well as the types of information it usually contains.
- The main users of the patient chart.
- The typical annotations and codes used by hospital workers on the patient chart.
- The physical location and level of mobility of the patient chart across the turn shift.
- Role of the patient chart on the activities carried out by operative nurses
- Other information artifacts used together with the patient chart.

At the end of the shadowing, semi-structured interviews complemented the observations and allowed us to inquiry informants with regards to the limits of the current version of the patient chart and implications of potential changes to it. During the study we also collected many samples of patient charts that served to conduct further analysis after the observations.

3 The Augmented Patient Chart

The analysis of the previous version of the patient chart resulted on a design that preserves the essential elements of that version, eliminates some of its problems, and provides new ways to capture data by taking advantage of the digital pen. The augmented patient chart system was developed and organized into two main components. The first component is the actual paper format that serves to implement the version of the patient chart using the digital pen and Anoto™ technology. The second component is the software system that processes the information captured with the digital pen. This section describes the characteristics of the system and discusses some of the design decisions made to implement it.

3.1 General Working Characteristics of the Digital Pen and Paper

Based on the Anoto™ technology, the system works with a digital pen with an embedded camera on it that takes pictures as the device moves across the surface while making annotations on the paper. The digital pen is just slightly bigger than a regular pen which means that people can use it without noticing its special characteristics. The paper contains a special printed dotted pattern that works with the camera to identify the exact position where the annotation is made on the paper. The area covered by the pattern used by Anoto™ allows creating a unique pattern for each physical form which allows designing applications consisting of multiple formats. For our particular case, we used a pattern for a letter-sized format (8.5 x 11 in).

3.2 System's Mode of Interaction and Operation

Our design of the augmented patient chart aimed at emphasizing the interaction with the paper format in order to preserve the natural work flow followed by nurses. Although some functions of the system (e.g. transferring data, or validating some entries) require the interaction with the computer-based interface, we defined the operation of the system in such a way that such interactions could be performed at the end of the working shifts. Figure 1 shows the new design proposed for the patient chart.

Following the original design, the frontal part of the augmented patient chart is organized into three main sections. At the top of the format a section is used to list patient's details such as name, age, gender, religion, and others. The middle section presents graphs for the cardiac pulse, internal and external temperature of the patient. Our design replaced manually generated graphs with a set of widgets to annotate accurate values used to plot graphs every time a new patient chart is created. The bottom section of the format is used to annotate (free format) additional information

about the conditions observed. The rear part of the format includes a section to annotate prescribed medications, general observations and specifying messages to other co-workers attending the patient.

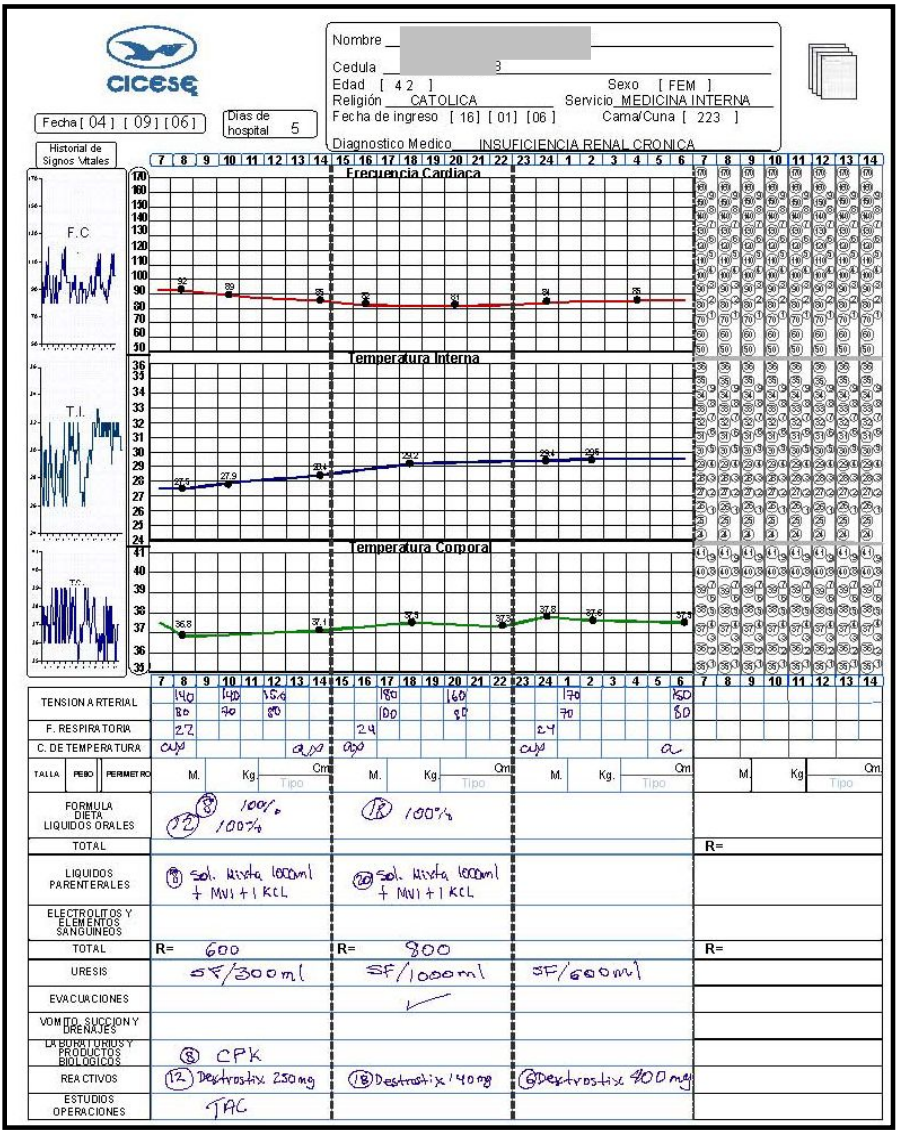


Fig. 1. The designed physical-digital patient chart (front)

The general operation of the system is detailed in the following lines: The nurse annotates information on the format as he does with the old version. Once he completes the annotations, he can leave the format and the digital pen and proceed

with his work. We assume that in general, it will be at the end of the work shift when the digital pen is placed in its base and then the information captured on the format is transferred to the system. Clearly, in some other situations immediate transfer might be necessary (as when a message has to be sent or medication has to be ordered) and the digital pen will have to be placed in its base right away. Whatever the case, once the pen is placed in its base, all information in memory is transferred to the system.

When the application detects that a particular action was requested (e.g. printing a new format or sending a message), the action is sent to a module that executes it.

The PC to which the digital pen is attached hosts a module named PLS (Paper Lookup Service) that detects and directs the action to a particular application, based on the pattern of the paper. The application then explores the areas of the paper that have been modified, obtains the information and processes it accordingly. Based on the annotation, the application executes actions such as updating records in the database, sending a message (through an electronic mail client), requesting confirmation of prescription orders or printing a new patient chart (including processing the information required to generate new graphs). Figure 2 shows the sequence diagram showing hot the components of the system interoperate.

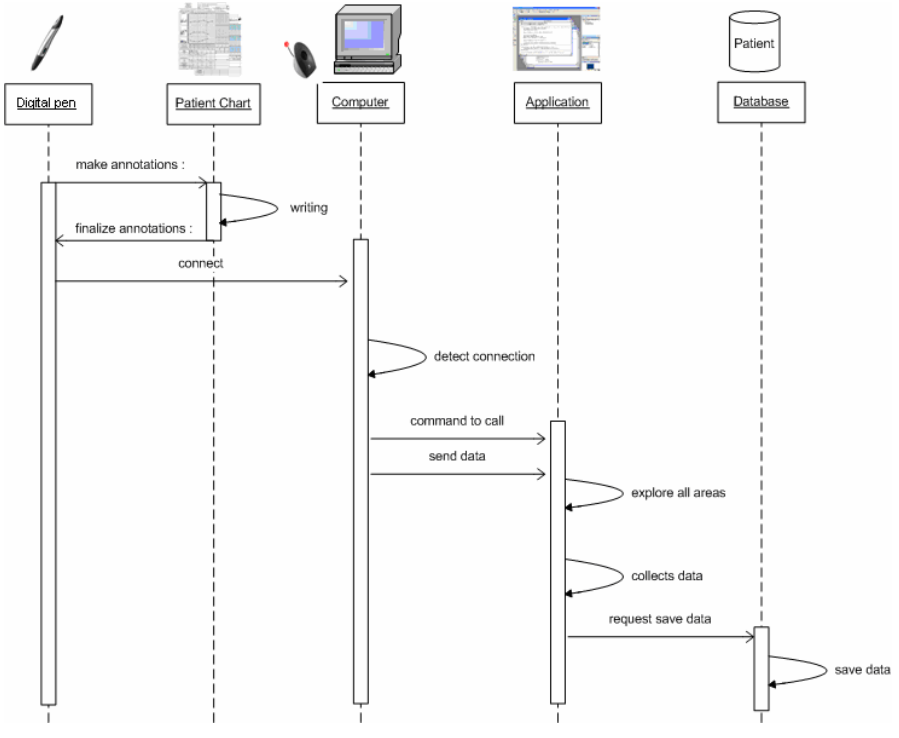


Fig. 2. Sequence diagram

In order to avoid errors, some annotations performed on the format have to be confirmed before an action takes place. Such is the case of the delicate task of ordering medication. Once the digital pen is placed in its base, the system detects

whether the medication section was modified. If new medications were annotated the system uses the captured images of the annotations and presents them on the computer screen. The user then has to manually scroll down among the list of medications, as well as other details such as dosage and periodicity, to select the one matching the annotation (see Figure 3).

Fig. 3. Confirmation of prescriptions

4 Evaluation

As a result of preliminary tests with users, the design of the system experienced a number of modifications until it eventually reached a level of maturity that allowed us to conduct a formal evaluation which results are presented in this section.

4.1 Purposes of the Evaluation and Methods

We aimed at conducting the evaluation to assess the advantages of the augmented patient chart format against the previous format used by nurses. Thus the evaluation of the system was guided by three main hypotheses:

H1. The time required to consult information using the new format is less than when using the traditional format.

We considered that the characteristics of the new format would produce that result because the information is structured in a better way, some annotations from the previous working shifts is printed with standard font and the values of graphs are more precisely plotted.

H2. The writing time using the new format is similar than when using the traditional format.

Our rational for this hypothesis was that the digital pen will not prove to be more challenging to use than the traditional one. Therefore, the time to write annotations would be very similar with both formats.

H3. The number of errors while either writing or finding information in the format will be less with the new format than with the traditional one.

We believed that this hypothesis would be true because the new format has a better structure, layout and leads to a better and more precise understanding of the information.

The three hypotheses above were tested in a controlled evaluation of the system. We decided to replicate as much as possible the conditions experienced by nurses while using the patient chart at the hospital. Thus, we designed our tasks and the evaluation setting with that purpose. The system was evaluated by 22 nurse students in Ensenada, Mexico (Figure 4). The nurse students were familiar with traditional patient charts because they have to spend practical hours at the local hospital as part of their education. After an initial training, the participants were asked to complete a set of representative tasks using both the traditional patient chart and the new version designed for the digital pen.

Fig. 4. Nurse students evaluating the system

The main working tasks that our evaluation aimed at exploring were: (1) reading of information on the format that was captured by another colleague during the previous work shift, and (2) annotation of information on the format as it would be done during the interaction with the patient; both tasks were performed by each participant using the traditional and the new patient chart formats. In order to minimize learning effects, we divided the participants and assigned them randomly to one of two groups. The first group (12 participants) started the evaluation using the digital format, and then used the traditional one. The second group (10 participants) started with the traditional format and then switched to the digital.

The evaluation was conducted in the following phases:

Phase 1. Using a laptop computer, each participant watched a video presentation describing the technology, as well as the set of tasks that they would be asked to complete. At the end of the presentation, the participant had the opportunity to clarify doubts or solve any question with the researcher.

Phase 2. During this phase the participants moved to another laptop computer and were given a copy of the augmented patient chart format. Standing in front of the computer, the participants have to respond a computer-based questionnaire (15

questions) on the computer by using the information contained on the patient chart format. This was the consulting information task.

Phase 3. After completing the previous questions, the participants moved to another computer and stood in front of it. They were asked to read information appearing on the computer screen that they have to use to annotate some sections of the patient chart format (7 tasks). After this, the nurses were asked to place the digital pen inside its base and then proceed to complete a confirmation of some of the information they gathered. This was the writing information task.

Phase 4. At this point, the participants received a copy of the traditional patient chart and moved to another computer to complete similar tasks as if they were dealing with another patient. Using the format they were asked to respond 14 questions appearing on the computer screen. It is important to clarify that the questions they have to respond during these phase were different to the ones they responded in a phase 2 since they were related to a different patient.

Phase 5. During this phase the participants move to the last computer station in the room and following the instruction on the computer they were asked to fill information on the format (7 tasks).

Phase 6. At the end of the evaluation, participants completed questionnaires based on TAM (Technology Adoption Model) [3] that served to explore their perspectives with regards to the adoption of our solution.

The first 12 participants followed the order described above. The next 10 participants completed the task in the following sequence: phase 1, phase 4, phase 5, phase 2, phase 3 and phase 6.

5 Results

The results of the evaluation indicate that using the new patient chart format participants experienced a significant reduction in the number of errors while reading information related to vital signs. Nine questions were asked to be responded by reading the graphs with vital sign information of the patient. On average, those volunteers using the traditional patient chart obtained significantly less correct answers (6.05 s.d. 1.35) compared to those using the new patient chart format (7.5, s.d. 1.79) ($t=3.864$, $p<0.001$). The other two sections of the format (evaluation and prescription of medicines) produced no significant differences in the number of errors when using either format.

We also found a significant increment of the accuracy while responding to the three tasks involving the annotating of data related to vital signs. On average the volunteers using the new patient chart format had more correct answers (5.18, s.d. 1.68) than those using the previous version of the format (3.68, s.d. 2.70) ($t = 2.28$, $p<0.05$). No advantage was observed for either version of the format while responding to tasks related to patient evaluation or prescription of medicines.

A trend towards less time spent while annotating data on the digital paper was observed. On average, those participants using the previous version of the format spend slightly more time (9.74 min. s.d. 3.01 min.) completing the tasks than those using the new digital format (9.10 min s.d. 2.52 min). Clearly, the difference is small, but still provides some evidence of the advantages of the new design.

Furthermore, based on the results from the TAM questionnaire, we found that the perceptions of participants with regards to the augmented patient chart's usefulness are very encouraging: they recognize the advantages of the solution, but more important, the new interface does not appear to be either challenging or fundamentally modifying the essence of the task performed with the artifact.

6 Conclusions

The use of physical paper artifacts working in combination with automated digital systems has been reported to be common not just in traditional office work contexts [6], but also as part of the practices of hospital and medical workers [2].

We proposed, implemented, and evaluated an augmented patient chart supporting a seamless integration of physical and digital artifacts for hospital work. Among its characteristics, the solution replaces manually generated graphs of patient's vital signs (cardiac pulse, internal and external temperature) with a set of widgets to annotate accurate values used to plot graphs every time a new patient chart is created. The system allows the free annotation of data such as medications that are validated when these are transferred to the hospital information systems. It also provides ways to annotate messages to co-workers and transfer them electronically. By combining widgets with pre-defined options and free annotations the system optimizes data capture without the need of explicit character recognition.

An evaluation of the patient chart system with 22 participants indicates a reduction in the number of errors while reading information, a significant increment on the accuracy while annotating data, as well as a trend towards less time spent while annotating data on the digital paper. The results of the TAM questionnaires indicate that users are inclined to adopt the system. Comments made by participants during and after the evaluation served to understand that users felt comfortable with the proposed design of the patient chart because it looks very similar to the original version and preserves the use of paper and pen as the main way to handle the information. These results allow us to argue that our design would be able to preserve the advantages associated with paper while increasing the availability of information and its trustworthiness.

Based on our understanding of hospital needs, we can identify a set of benefits resulting from using the type of technology we are proposing. Firstly, the level of precision and clarity is clearly increased which can, in principle, contribute towards the reduction of errors while managing information. Secondly, as a result of the architecture of the system and the way it manages the data, the digital format can be integrated and operate with other hospital information systems (e.g. Electronic Patient Record, Clinical Decision Support Systems, Computerized Physician Order Entry Systems, etc) without major modifications. Thirdly, because data is preserved in digital and physical versions it allows recreating the formats (and hand written annotations) in case of damage or lost. Finally, because the information is stored in databases it can be consulted without retrieving the physical artifact.

As part of our future work we plan to extend the functionality of the system and conduct an in situ extended trial of the system at a local hospital.

Acknowledgments. We wish to express our gratitude to all the volunteer nurse students that took part during the different stages of the design of the augmented patient chart, as well as to Raul Romero for his support organizing the evaluation, and to Monica Tentori, Elisa Moran, Leonardo Galicia, Daira Sanchez and Mario García for their support during the evaluation of the system.

References

1. Bardram, J., Bossen, C.: Mobility Work: The Spatial Dimension of Collaboration at a Hospital. Journal of Computer Supported Cooperative Work (Springer, Netherlands) 14(2), 131–160 (2005)
2. Bossen, C.: The Parameters of Common Information Spaces: the Heterogenity of Cooperative Work at a Hospital Ward. In: Proceedings of ACM Conf. on Computer Supported Cooperative Work, CSCW, pp. 176–185. ACM Press, New York (2002)
3. Davis, F.: Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly 13(3), 319–339 (1989)
4. Kohn, L., Corrigan, J., Donaldson, M. (eds.): To Err is Human: Building a Safer Health System. Committee on Quality of Health Care in America. Institute of Medicine. Washington, DC, USA (1999)
5. Theureau, J., Filippi, G.: Analysing cooperative work in an urban traffic control room for the design of a coordination support system. In: Workplace Studies: Recovering Work Practice and Informing System Design, pp. 68–91. Cambridge Press, Cambridge (2000)
6. Sellen, A.J., Harper Richard, R.: The Myth of the Paperless Office. 56 MIT Press, Cambridge, Mass (2000)
7. Silva, J., Zamarripa, M., Strayer, P., Favela, J., Gonzalez, V.: Empirical Evaluation of a Mobile Application for Assisting Physicians in Creating Medical Notes. In: the Proceedings of the 12th Americas Conference on Information Systems. Acapulco, Guerrero, Mexico (2006)

Author Index

Abascal, Julio 3
Adams, Ray 473
Ahamed, Sheikh Iqbal 833
Aizpurua, Amaia 3
Albert, Dietrich 647
Allely, Eric 979
Ardito, C. 530
Argyropoulos, Vassilis 205
Arrue, Myriam 3
Asai, Daisuke 185
Asakawa, Chieko 11, 98
Asano, Yoko 185, 961
Auinger, Andreas 481

Bain, Keith 214, 446
Bardzell, Jeffrey 860
Behringer, Reinhold 520
Bennett, Steve 224
Bevan, Nigel 20
Biterman, Noemi 853
Bitterman, Haim 853
Blevis, Shunying 860
Bocconi, Stefania 491
Bogren, Linda 269
Boticario, Jesús G. 796
Bouhli, Maria 941
Boussakta, Said 874
Brajnik, Giorgio 501
Brezin, Jonathan P. 133
Busuoli, M. 511
Buzzi, Marina 687

Carriço, Luís 259
Chautard, David 677
Chen, Min-Ting 279
Chen, Lih-Yuan 234
Chen, Ming-Chung 234, 242, 371, 462
Chen, Su-Ying 242
Chêne, Denis 598
Chiang, Chih-Cheng 279
Chiang, Chun-Han 234
Chitnis, S. 309
Cho, Joo Eun 64
Choi, Doojin 55
Christian, Johannes 520
Chu, Chi Nung 869
Cko, Chien-Chuan 234, 242
Claridge, Nigel 20
Clure, Gary 716
Conlan, Owen 647
Costabile, M.F. 530
Crayne, Susan 133
Ćukušić, Maja 617

de Ruiter, J.P. 251
Debevc, Matjaž 539, 549
Diez-Diaz, Fidel 28
Ding, Jiarong 697
Dini, Silvia 491
Duarte, Carlos 259

Ebner, Martin 559
Enfield, N.J. 251
Engelen, Jan 160
Exadaktylos, George 823

Fabián, V. 511
Fallman, Daniel 269
Fan, Jianmiao 389
Fan, Yongkai 569
Favela, Jesus 1006
Fei, Yu-Ming 279
Ferlino, Lucia 491
Fernández del Viso, Alicia 796
Fink, Josef 813
Finlayson, J. Louise 289
Fjellstrom, Oskar 269
Fourli-Kartsouni, Florendia 299
Fu, Xiaolong 569

Galáz, Iván 786
Gallelli, T. 511
Gappa, Henrike 124
Glavinic, Vlado 578
Go, Christian Anthony L. 588
Goggin, Ashley 349
Gonzalez, Victor M. 1006
Gonzalez-Rodriguez, Martin 28
Good, Alice 34

Gopal, D. 309
Gouy-Pailler, Cédric 598
Grammenos, Dimitris 607
Granić, Andrina 617, 627
Grigoriadou, Maria 319
Guo, H. 309
Gupta, Gopal 309, 329
Gutiérrez y Restrepo, Emmanuelle 796

Hadjiefthymiades, Stathes 406
Haluzík, M. 511
Handa, Takuya 45, 88
Hanson, Vicki L. 133
Hasebe, Harumi 416
Herbst, Ilona 738
Hewitt, Jill 224
Hines, Jason 214
Hiruma, Nobuyuki 45
Ho, Chun-Min 380
Holzinger, Andreas 520, 559, 728, 813
Hone, Kate S. 932
Hsu, Kuang-Chieh 453
Huang, Yu Ting 869
Hwang, Sheue-Ling 997
Hyun, Joonho 55

Ikechi, Oluchi 932
Ishihara, Tatsuya 426
Ito, Takayuki 45
Itoh, Kazunori 806
Itoh, Takashi 11

Jain, Aanchal 329
Jerrams-Smith, Jenny 34
Johansson, Lars 269

Kanatsugu, Yasuaki 45, 88
Kang, Guixia 874, 883
Karim, Shuaib 637
Karshmer, A. 309
Karvonen, Kristiina 893
Katsoulis, Philippos 205
Kickmeier-Rust, Michael D. 647
Kim, Byung Geun 73
Kim, Chang-Soo 903
Kim, Jong-Woo 903
Kim, Jong Gon 73
Kim, Ju-Yeon 903
Kim, Sukil 55
King, Neil 124

Kjeldsen, Rick 339
Koch, Johannes 160, 177
Kokol, Peter 549
Kondratova, Irina 349
Kosec, Primož 539
Kourbani, Vassiliki 657
Kouroupetroglou, Christos 358
Kouroupetroglou, Georgios 205, 299, 436
Krieger, Horst 520
Kuber, Ravi 913
Kuebler, Ronald C. 368
Kuebler, Scott H. 368
Kuljis, Jasna 667

Laff, Mark R. 133
Lanzilotti, R. 530
Latif, Khalid 637
Lee, Seongil 64, 73
Lee, Won-Hyung 588
Leikas, Jaana 923
Lenz, Franz-Josef 728
Lepicard, Guillaume 677
Leporini, Barbara 687
Lerner, Eyal 853
Leung, Cherng-Yee 453
Li, Christina 79
Li, Jason 697
Li, Shanghong 874
Lin, Huei-Tang 453
Lin, Jun 569
Lin, Ya-Li 380
Lin, Yun-Lung 371
Lin Chiuhsiang, Joe 279
Lindbergh, Kent 269
Lines, Lorna 667, 932
Lingras, Pawan 214
Lisney, Eleanor 79
Liu, Da 883
Liu, Hunszu 399
Liu, Songtao 79
Liu, Yong 389
Lyon, Caroline 224

Manitsaris, Athanasios 358
Margetis, George 941
Martinoli, Cristina 491
Masthoff, Judith 289
Matsumura, Kinji 45, 88
Matsunami, Haruhito 951

Maurer, Hermann 559
McAllister, Graham 913
Mellish, Chris 289
Mellor, Barry 224
Miura, Taiki 116
Miyashita, Hisashi 11, 98, 167
Moraveji, Neema 697
Movahedi, Jafar 108
Musioł, Teresa 707
Muta, Hidemasa 142

Nakić, Jelena 627
Naoshi, Tabuchi 167
Nathanail, Spyros 406
Nazari Shirehjini, Ali A. 108
Newman, Dianna L. 716
Nilbrink, Fredrik 269
Nischelwitzer, Alexander 728
Nishizawa, Yosoko 416
Nordbrock, Gaby 124
Novák, D. 511
Ntoa, Stavroula 941

O'Kelley, Patrick 697
Ohko, Takuya 142
Okada, Hidehiko 116
Ono, Kouichi 167
Ortner, Daniela 160
Ott, Michela 491
Ouyang, Yue 883

Papadopoulos, Konstantinos 205
Peirce, Neil 647
Pérez de la Cámara, Santiago 796
Petrie, Helen 20, 124
Petropoulos, Yiannis 657
Pohl, Margit 738
Povalej, Petra 549
Pozzi, Francesca 748
Procter, Margaret 756
Pucheu, Elian 677

Qin, Yumei 214

Rankin, Jeff 349
Rebate Sánchez, Carlos 796
Reichl, Franz 738
Richards, John T. 133
Riess, Judith 979
Rigolleau, Benoît 677
Rose, J. Barbara 766
Roselli, T. 530
Rosic, Marko 578
Rossano, V. 530
Rubin, Bella 776

Sagata, Yoshie 961
Saito, Shin 167
Sakai, Tadahiro 45, 88
Sakairi, Takashi 142
Salampasis, Michail 358
Sánchez, Jaime 786
Santos, Olga C. 796
Sarid, Helen 776
Sato, Daisuke 98
Savidis, Anthony 607
Sawicka, Agata 150
Schwarz, Daniel 647
Searle, Gig 728
Shimizu, Michio 806
Shinya, Atsushi 970
Slavakis, Kostas 299
Snaprud, Mikael 150, 177
Sorensen, Barbara 979
Stary, Christian 481
Štěpánková, O. 511
Stephanidis, Constantine 607, 941
Stickel, Christian 813
Stjepanovič, Zoran 539, 549
Strobbe, Christophe 160, 177
Sugimoto, Masahiko 806
Sui, Yangyi 569
sun, Tianze 569
Swart, Calvin B. 133

Takagi, Hironobu 11, 98, 426
Tateishi, Takaaki 167
Theodoridis, Sergios 299
Tjoa, A. Min 637
Tolias, Dimitris 657, 823
Tsaganou, Grammatiki 319
Tsetsos, Vassileios 406
Tsonos, Dimitrios 436
Tu, Yung-Hsiang 453
Tyagi, Vaibhav 833

Ujma-Wąsowicz, Katarzyna 707

Velasco, Carlos A. 124, 160, 177
Vella, Frédéric 677

Velleman, Eric 177
Verlič, Mateja 539, 549
Verpoorten, Dominique 647
Vidal, Sylvie 598
Vidau, Agueda 28
Vigo, Markel 3
Vigouroux, Nadine 677
Vlachogiannis, Evangelos 160

Wakizaka, Yoshinori 416
Wald, Mike 446
Wang, Huaqing 883
Wang, Hwa-Pey 371
Wang, Q. 309
Watanabe, Masahiro 185, 961
Weber, Gerhard 193
Weimann, Kurt 193
Wiltner, Sylvia 738
Woolf, Suze 697
Wroblewski, Nancy 860
Wu, Chih-Fu 453
Wu, Siew-Rong 843
Wu, Ting-Fang 242, 462
Wu, Ya-Ping 371

Xydas, Gerasimos 205, 436

Yamamoto, Takayuki 416
Yamaoka, A. Toshiki 990
Yang, Yi-Fan 997
Yeh, Chih-Ching 242
Yeh, Yao-Ming 371, 869
Yoshida, Mayuko 990
Yu, Wai 913

Zamarripa, Myrna S. 1006
Zelic, Marija 578
Zhang, Li 874
Zhang, Ping 874, 883
Zijp-Rouzier, Sophie 598