AMERICAN UNIVERSITY STUDIES

The Concept of Logical Consequence

An Introduction to Philosophical Logic

MATTHEW W. McKEON

The Concept of Logical Consequence

AMERICAN UNIVERSITY STUDIES

SERIES V
PHILOSOPHY

VOL. 207

PETER LANG
New York • Washington, D.C./Baltimore • Bern
Frankfurt • Berlin • Brussels • Vienna • Oxford

Matthew W. McKeon

The Concept of Logical Consequence

An Introduction to Philosophical Logic

PETER LANG
New York • Washington, D.C./Baltimore • Bern
Frankfurt • Berlin • Brussels • Vienna • Oxford

Library of Congress Cataloging-in-Publication Data

McKeon, Matthew W.
The concept of logical consequence: an introduction
to philosophical logic / Matthew W. McKeon.
p. cm. — (American university studies. V, Philosophy; v. 207)
Includes bibliographical references.
1. Logic. I. Title.
BC71.M43 160—dc22 2009018353
ISBN 978-1-4331-0645-3
ISSN 0739-6392

Bibliographic information published by **Die Deutsche Nationalbibliothek**.
Die Deutsche Nationalbibliothek lists this publication in the "Deutsche Nationalbibliografie"; detailed bibliographic data is available on the Internet at http://dnb.d-nb.de/.

The paper in this book meets the guidelines for permanence and durability of the Committee on Production Guidelines for Book Longevity of the Council of Library Resources.

29 Broadway, 18th floor, New York, NY 10006
www.peterlang.com

Printed in Germany

For Beth

Table of Contents

Chapter 1
Introduction

This book is an inquiry into the concept of logical consequence, arguably the central concept of logic. We take logical consequence to be a relation between a given set of sentences and the sentences that logically follow. One sentence is said to be a logical consequence of a set of sentences, if and only if, in virtue of logic alone, it is impossible for the sentences in the set to be all true without the other sentence being true as well. The central question to be investigated here is: what conditions must be met in order for a sentence to be a logical consequence of others?

One historically significant answer derives from the work of Alfred Tarski, one of the greatest logicians of the twentieth century. In Chapter 2, we distinguish features of the ordinary, informal concept of logical consequence using some of Tarski's work, particularly his seminal (1936) paper on logical consequence. Here Tarski uses his observations of the salient features of what he calls the common concept of logical consequence to guide his theoretical development of it. We shall develop Tarski's observations of the criteria by which we intuitively judge what follows from what, and which Tarski thinks must be reflected in any theory of logical consequence.

After presenting his theoretical definition of logical consequence, which is the forerunner of the modern, model-theoretic definition, Tarski asserts in his (1936) paper that it reflects the salient features of the common concept of logical consequence. This assertion is not obvious, and Tarski defends it nowhere in his published writings. This raises the particular issues of whether Tarski's informal characterization of the common concept of logical consequence is correct, and whether it is reflected in his theoretical definition. The more general issues raised are: how do we justify a theoretical definition of logical consequence? What role should the informal concept play?

We shall answer these questions with respect to the model-theoretic and the deductive-theoretic characterizations of logical consequence for first-order languages. They represent two major theoretical approaches to making the common concept of logical consequence more precise. Chapter 2 shall motivate both approaches by considering them as natural developments of the ordinary, informal characterization. This shall set the context for our critical evaluation of these two approaches to characterizing logical consequence. After introducing some set-theoretic concepts used in the book and a simple first-order (extensional) language M in Chapter 3, (classical) logical consequence shall be defined for M model-theoretically in Chapter 4, and deductive-theoretically (a natural deduction system N is given) in Chapter 5. I account for their status as definitions, and sketch how they work in determining what follows from what. Also, there are accounts of what models and deductive apparatuses are, and, what, exactly, they represent when used to fix the logical consequence relation.

Both Chapters 4 and 5 consider methodological criticism of the model-theoretic and deductive-theoretic approaches. In particular, we consider the adequacy of models and deductive apparatuses as tools for defining logical consequence, and these considerations are used to answer the two questions posed above: how do we justify a theoretical definition of logical consequence? What role should the informal concept play? Also, in Chapters 4 and 5, there is some criticism of classical logic. Both types of criticism (methodological and logical) not only motivate consideration of alternative logics, but also suggest revisions to the Tarskian understanding of the informal concept of logical consequence introduced in Chapter 2.

While most logicians accept the model-theoretic and deductive-theoretic characterizations of logical consequence for extensional languages, there is less agreement on the pre-theoretic notion these technical definitions are supposed to represent, and little discussion about whether they actually do represent it adequately. Almost all of the formal logic textbooks written for the book's intended audience give an ordinary, informal characterization of logical consequence either in the introduction or at the beginning of the first chapter. Unfortunately, the informal characterization of logical consequence typically amounts to a mere sketch which is either insufficient for clarifying the status of the technical characterizations that follow or conflicts with them. The book's focus on the concept of logical consequence, its introductory manner of presentation, and it's monograph-length, make it ideal for the intended audience as a means for clarifying the status and aims of the technical characterizations of logical consequence, and for highlighting their relationship to the informal concept of logical consequence which motivates them. This enhances understanding of not only the status of the model-theoretic and deductive-theoretic characterizations of logical consequence, but also deepens our understanding of criteria for evaluating them.

The book's intended audience matches the audiences of other introductions to philosophical logic such as Haack (1978), Sainsbury (1991), and

Read (1995). Like these classics, this book is written at a level that makes it beneficial to advanced undergraduates with exposure to introductory formal logic, graduate students, and professional philosophers planning to self-educate themselves about the philosophy of logical consequence and for whom this book is only a first step. What distinguishes this book is its approach to thinking about logical consequence. It is tightly organized around the informal concept of logical consequence, and its relationship to the more technical model-theoretic and deductive-theoretic characterizations. I am unaware of any introduction to philosophical logic devoted to motivating the technical characterizations of logical consequence by appealing to the informal concept of logical consequence, and evaluating the former in terms of how successfully they capture the central features of the latter. As with the above three books when first published, the freshness of this book's approach to studying logical consequence and its engagement with themes in the recent literature should make it of interest to specialists working in the philosophy of logic.

The goal of realizing the envisioned length of the book has, of course, had an expository impact. In order to forgo lengthy exegetical analysis, ideas and arguments from the literature are typically presented in a distilled form. Also, since references and discussion are confined to the main text, there are no footnotes. More importantly, I have been careful in choosing where to be argumentative (as in my defense of the Tarksian model-theoretic characterization of logical consequence against criticism) and where to remain agnostic (as with respect to the issue of the nature of a logical constant and whether the meanings of logical constants should be identified with their truth-conditional properties or there inferential properties). I've chosen to be argumentative in those places where I believe that I have the space to be persuasive. In many places where the discussion is more expository and less argumentative, I have developed topics to the point that satisfies the stated goals of the relevant section. I certainly realize that on pretty much every topic covered in the book, much more can be usefully said. I have at points in the text provided the reader with references that extend the book's discussion in various ways.

There are references to Tarski's work throughout the book's discussion of logical consequence. While I do not believe that my reading of Tarski is controversial, the desire not to lengthen the book prohibits defense of my interpretation of Tarski. To be clear, this is not a book about Tarski. Rather, some of Tarski's writings are used as a platform for the book's discussion of logical consequence. The thoughts of other logicians such as Dummett, Gentzen, and Frege, are also used towards this end. The trajectory of the discussion is squarely aimed at the model-theoretic and deductive-theoretic approaches to logical consequence, and their relationship to the informal concept of logical consequence. Even though the focus of the book is on logical consequence, it is studied in a way that allows it to serve as an introduction to philosophical logic. Its emphasis on the informal concept of logical consequence and its relationship to the more technical model-theoretic and deductive-theoretic approaches highlights and sharpens in a unique way other issues central in the philosophy of logic such as the nature of logic, logical constants, and logical necessity. Introducing issues in the philosophy of logic from the perspective of a study of logical consequence will illustrate how these issues are related, and why they are significant for understanding logical consequence.

Chapter 2
The Concept of Logical Consequence

Tarski's Characterization of the Common Concept of Logical Consequence

Tarski begins his article, "On the Concept of Logical Consequence," by noting a challenge confronting the project of making precise the common concept of logical consequence.

> The concept of *logical consequence* is one of those whose introduction into a field of strict formal investigation was not a matter of arbitrary decision on the part of this or that investigator; in defining this concept efforts were made to adhere to the common usage of the language of everyday life. But these efforts have been confronted with the difficulties which usually present themselves in such cases. With respect to the clarity of its content the common concept of consequence is in no way superior to other concepts of everyday language. Its extension is not sharply bounded and its usage fluctuates. Any attempt to bring into harmony all possible vague, sometimes contradictory, tendencies which are connected with the use of this concept, is certainly doomed to failure. We must reconcile ourselves from the start to the fact that every precise definition of this concept will show arbitrary features to a greater or less degree. ((1936), p. 409)

Not every feature of a precise definition of logical consequence will be reflected in the common concept of logical consequence, and we should not expect any precise definition to reflect all of its features. Nevertheless, despite its vagueness, Tarski believes that there are identifiable, essential features of the common concept of logical consequence.

> …consider any class K of sentences and a sentence X which follows from this class. From an intuitive standpoint, it can never happen that both the class K consists of only true sentences and the sentence X is false. Moreover, since we are concerned here with the concept of logical, i.e., formal consequence, and thus with a relation which is to be uniquely determined by the form of the sentences between which it holds, this relation cannot be influenced in any way by empirical knowledge, and in particular by knowledge of the objects to which the sentence X or the sentences of class K refer. The consequence relation cannot be affected by replacing designations of the objects referred to in these sentences by the designations of any other objects. (1936, pp.414-415)

According to Tarski, the logical consequence relation is (1) necessary, (2) formal, and (3) not influenced by empirical knowledge. We now elaborate on (1)-(3).

The logical consequence relation has a modal element

Tarski countenances an implicit modal notion in the common concept of logical consequence. If X is a logical consequence of K, then not only is it the case that not all of the sentences of K are true and X is false, but it can *never* happen that both the class K consists of only true sentences and the sentence X is false. That is, X logically follows from K only if it is necessarily true that if all the sentences in K are true, then X is true, i.e., it is not possible for all the K-sentences to be true with X false. For example, the supposition that *All West High School students are football fans* and that *Kelly is not a West High School student* does not rule out the possibility that Kelly is a football fan. Hence, the sentences *All West High School students are football fans* and *Kelly is not a West High School student* do not entail *Kelly is not a football fan*, even if she, in fact, isn't a football fan. Also, *Most of Kelly's male classmates are football fans* does not entail *Most of Kelly's classmates are football fans*. What if the majority of Kelly's class is composed of females who are not fond of football?

The sentences *Kelly is not both at home and at work* and *Kelly is at home* jointly imply that *Kelly is not at work*. Note that it doesn't seem possible for the first two sentences to be true and *Kelly is not at work* false. But it is hard to see what this comes to without further clarification of the relevant notion of possibility. For example, consider the following pairs of sentences.

Kelly is a female.
Kelly is not the US President.

Kelly kissed her sister at 2:00pm.
2:00pm is not a time during which Kelly and her sister were ten miles apart.

There is a chimp in Paige's house.
There is a primate in Paige's house.

Ten is greater than nine.
Ten is a prime number.

For each pair, there is a sense in which it is not possible for the first to be true and the second false. At the very least, an account of logical consequence must distinguish logical possibility from other types of possibility. Should truths about physical laws, US political history, zoology, and mathematics constrain what we take to be possible in determining whether or not the first sentence of each pair could logically be true with the second sentence false? If not, then this seems to mystify logical possibility (e.g., how could ten be a prime number?). Given that I know that Barack Obama is US President and that he is not a female named *Kelly*, isn't it inconsistent for me to grant the logical possibility of the truth of *Kelly is a female* and the falsity of *Kelly is not the US President*? Or should I ignore my present state of knowledge in considering what is logically possible? Tarski does not derive clear notions of the logical modalities (i.e., logical necessity and possibility) from the common concept of logical consequence. Perhaps there is none to be had, and we should seek the help of a proper theoretical development in clarifying these modal notions. With this end in mind, let's turn to the other features of logical consequence highlighted by Tarski, starting with the formality criterion of logical consequence.

The logical consequence relation is formal

Tarski observes that logical consequence is a formal consequence relation. And he tells us that a formal consequence relation is a consequence relation that is uniquely determined by the sentential forms of the sentences between which it holds. Consider the following pair of sentences.

1. Some children are both lawyers and peacemakers
2. Some children are peacemakers

Intuitively, (2) is a logical consequence of (1). It appears that this fact does not turn on the subject matter of the sentences. Replace 'children', 'lawyers', and 'peacemakers' in (1) and (2) with the variables *S*, *M*, and *P* to get the following.

1'. Some S are both M and P
2'. Some S are P

(1') and (2') are sentential forms (*sentential functions*, in Tarski's terminology) of (1) and (2), respectively. Note that there is no interpretation of *S*, *M*, and *P* according to which the sentence that results from (1') is true and the resulting instance of (2') is false. Hence, (2) is a formal consequence of (1), and on each meaningful interpretation of *S*, *M*, and *P* the resulting (2') is a formal consequence of the sentence that results from (1') (e.g., *some clowns are sad* is a formal consequence of *some clowns are lonely and sad*). Tarski's observation is that, relative to a language L, for any sentence X and class K of sentences, X is a logical consequence of K only if X is a formal consequence of K. The formality criterion of logical consequence can work in explaining why one sentence doesn't entail another in cases where it seems impossible for the first to be true and the second false.

For example, to think that (3) *Ten is a prime number* does not entail (4) *Ten is greater than nine* does not require one to think that ten could be a prime number and less than or equal to nine, which is a good thing since it is hard to see how this is possible. Rather, we take (3') *a* is a P and (4') *a* is R *b* to be the forms of (3) and (4) and note that there are interpretations of 'a', 'b', 'P', and 'R' according to which the first is true and the second false (e.g., let 'a' and 'b' name the numbers two and ten, respectively, and let 'P' mean *prime number*, and 'R' *greater than*). Note that the claim here is not that formality is sufficient for a consequence relation to qualify as logical, but only that it is a necessary condition. I now elaborate on this last point by saying a little more about sentential forms and formal consequence.

Distinguishing between a term of a sentence replaced with a variable and one held constant determines a form of the sentence. In *Some children are both lawyers and peacemakers* we may replace 'Some' with a variable and treat all the other terms as constant. Then

1". D children are both lawyers and peacemakers

is a form of (1), and each sentence generated by assigning a meaning to *D* shares this form with (1). For example, the following three sentences are instances of (1"), produced by interpreting *D* as 'No', 'Many', and 'Few'.

No children are both lawyers and peacemakers
Many children are both lawyers and peacemakers
Few children are both lawyers and peacemakers

Whether X is a formal consequence of K then turns on a prior selection of terms as constant and others uniformly replaced with variables. Relative to such a determination, X is a formal consequence of K if and only if (hereafter we abbreviate 'if and only if' to 'iff') there is no *interpretation* of the variables according to which each of the K-sentences are true and X is false.

So, taking all the terms, except for ‘Some’, in (1) *Some children are both philosophers and peacemakers* and in (2) *Some children are peacemakers* as constants makes the following forms of (1) and (2).

1". D children are both lawyers and peacemakers
2". D children are peacemakers

Relative to this selection, (2) is not a formal consequence of (1), because replacing ‘D’ with ‘No’ yields a true instance of (1") and a false instance of (2").

Tarski’s observation in his (1936) article that if we treat all expressions of a language L as logical constants, then logical consequence between L sentences collapses into what he calls material consequence is an obstacle to developing Tarski’s view of logical necessity and possibility. A sentence X is a material consequence of a class K of sentences if either X is true or at least one of the K-sentences is false. The reduction of material consequence to logical consequence relative to treating all expressions as logical is problematic, because material consequence does not reflect logical necessity. That “X is either true or at least one of the K-sentences” is false does not capture “it is logically necessary that if all the K-sentences are true then X is true.” For example, ‘Hillary Clinton is not Vice-President’ is true or ‘Barack Obama is US President’ is false, but it doesn’t seem necessarily the case that if ‘Barack Obama is US President’ is true, then ‘Hillary Clinton is not Vice-President’ is true. Tarski may here have identified formal consequence with logical consequence (which he explicitly does on p.414, (1936)). It is certainly true that on the above scenario formal consequence collapses into material consequence, but as we shall highlight below, a formal consequence relation is a logical consequence relation only if the expressions held constant are logical. Unfortunately, Tarski does not explain in his (1936) article what qualifies as a logical expression, and so provides no philosophical rationale for treating all of a language’s expressions as logical.

Tarski briefly considers a characterization of formal consequence according to which a sentence X is a formal consequence of class K of sentences iff there is no *uniform substitution* of the variables with expressions available from the relevant language that results in each of the K-sentences being true and X false ((1936), p.417). Tarski thinks that this notion of formal consequence cannot be used to make precise the common concept of logical consequence, because the former deviates from the later with respect to impoverished languages. For example, we regard (3') and (4') as the sentential forms of (3) and (4), respectively.

3. Ten is a prime number
4. Ten is greater than nine

3'. *a* is a P
4'. *a* is R *b*

Suppose that the only expressions of the relevant language L are those in (3) and (4), i.e., 'ten', 'nine', 'prime number', and 'greater than'. Obviously, there is no uniform substitution of the variables in (3') and (4') with expressions from L that results in a true instance of (3') and a false instance of (4'). So, (4) is a formal consequence, thus conceived, of (3), and, arguably, it can never happen that (3) and the negation of (4) are true. Nevertheless, it doesn't seem like (4) should be a logical consequence of (3). As noted above, there are interpretations of 'a', 'b', 'P', and 'R' according to which (3') is true and (4') is false, which seem relevant to showing that (4) is not a logical consequence of (3). The extension of the common concept of logical consequence in a language should not essentially depend on the richness of the language, since its expressive resources seems to be a contingent matter.

Tarski, in effect, notes that if a language has names of all possible objects, then a formal consequence relation that appeals to uniform substitution is equivalent to one that appeals to interpretations. The appeal to uniform substitution would be adequate to the task of characterizing the common concept of logical consequence as applied to sentences of the relevant language. However, as Tarski remarks, no language contains names of all possible objects. At any rate, the consideration at the end of the previous paragraph are reason enough to maintain that whether one sentence is a logical consequence of others shouldn't depend on the richness of relevant language.

We now return to the characterization of formal consequence in terms of interpretations. Recall that on this on this characterization a sentence X from a language L is a formal consequence of a class K of sentences, because, relative to a selection of expressions from L as constant and others uniformly replaced with variables, there is no interpretation of the variables according to which each of the K are true and X is false. It is easy to show that formal consequence—in this sense—is not sufficient for logical consequence.

Consider the following pair of sentences.

5. Kelly is female
6. Kelly is not US President

(6) is a formal consequence of (5) relative to just replacing 'Kelly' with a variable. Given current U.S. political history, there is no individual whose name yields a true (5) and a false (6) when it replaces 'Kelly'. This is not, however, sufficient for seeing (6) as a logical consequence of (5). There are two ways of thinking about why, a metaphysical consideration and an epistemological one. First the metaphysical consideration: it seems possible for (5) to be true and (6) false. The course of U.S. political history could have turned out differently. One might think that the current US President could—logically—have been a female named, say, 'Hillary'. Using 'Hillary' as a replacement for 'Kelly' would yield in that situation a true (5) and a false (6). Also, it seems possible that in the future there will be a female US President. This suggests that a formal consequence relation determined by treating 'female' and 'US President' as constants fails to reflect the modal notion inherent in the common concept of logical consequence. In order for a formal consequence relation from K to X to qualify as a logical consequence relation it has to be the case that it is *necessary* that there is no interpretation of the variables in K and X according to which the K-sentences are true and X is false. The epistemological consideration is that one might think that knowledge that X follows logically from K should not essentially depend on being justified by experience of extra-linguistic states of affairs. Clearly, the determination that (6) follows formally from (5) essentially turns on empirical knowledge, specifically knowledge about the current political situation in the US.

Of course, the point that formal consequence is not sufficient for logical consequence does not rule out the existence of constants that determine a formal consequence relation invulnerable to versions of the above metaphysical and epistemological considerations. Before considering what qualifies as a logical constant, we turn to the final highlight of Tarski's rendition of the intuitive concept of logical consequence: that knowledge of logical consequence cannot be empirical.

The logical consequence relation is *a priori*

Tarski says that by virtue of being formal, knowledge that X follows logically from K cannot be affected by knowledge of the objects that X and the sentences of K are about. Hence, there is an epistemic feature of the logical consequence relation: our knowledge that X is a logical consequence of K cannot be influenced by empirical knowledge. However, as noted above, formality by itself does not insure that knowledge of the extension of a consequence relation is unaffected by knowledge of the objects that X and the sentences of K are about. Here we deviate from Tarski and view the alleged epistemic feature of logical consequence as independent of formality. We characterize empirical knowledge in two steps as follows. First, *a priori* knowledge is knowledge "whose truth, given an understanding of the terms involved, is ascertainable by a procedure which makes no reference to experience" (Hamlyn (1967), p. 141). Empirical, or *a posteriori*, knowledge is knowledge that is not *a priori*, i.e., knowledge whose validation necessitates a procedure that does make reference to experience. Knowledge that

2+2=4,
if I am conscious, then I exist, and
the shortest distance between two points in a plane is a straight line,

is *a priori*. While knowledge that

Barack is a Democrat,
if Kelly is female, then she is not the US President, and
the Pope is older than the US President,

is empirical knowledge. We read Tarski as saying that a consequence relation is logical only if that something falls in its extension is knowable *a priori*, i.e., only if the relation is *a priori*. We may call this the *a prioricity* criterion of logical consequence.

Knowledge of physical laws, a determinant in people's observed sizes, is not *a priori* and such knowledge is required to know that there is no interpretation of *k*, *h*, and *t* according to which (7) is true and (8) false.

7. k kissed h at time t
8. t is not a time during which k and h are ten miles apart

So (8) cannot be a logical consequence of (7). However, my knowledge that *Kelly is not Paige's friend* follows from *Kelly is taller than Paige's only friend* is *a priori* since, arguably, I know *a priori* that *nobody is taller than herself*.

In an English translation of the Polish version of Tarski's paper on logical consequence, Tarski offers a different formulation of the epistemological feature of logical consequence: "…[the logical consequence relation] cannot depend on our knowledge of the external world, in particular on our knowledge of the objects spoken about in [the relevant sentences]" (Stroińska and Hitchcock (2002), p. 184). Obviously, the *a prioricity* criterion of logical consequence rules out that logical consequence depends on the obtaining of extra-linguistic states of affairs only when joined with the thesis that all knowledge of extra-linguistic states of affairs is *a posteriori*. This empiricist thesis along with the *a prioricity* criterion of logical consequence, forces the idea that logical knowledge is linguistic knowledge, and, therefore, that a sentence X is a logical consequence of a class K of sentences is entirely a matter of the meanings of these sentences.

The *a prioricity* criterion is not uncontroversial. Indeed, Tarski later changed his views on the relationship between logical and empirical knowledge. In a letter written in 1944, Tarski demurs from the *a prioricity* criterion, remarking that, "I would be inclined to believe (following J.S. Mill) that logical and mathematical truths do not differ in their origin from empirical truths—both are results of accumulated experience" ((1987/1944), p. 31). And then later, "I think that I am ready to reject certain logical premises (axioms) of our science in exactly the same circumstances in which I am ready to reject empirical premises (e.g., physical hypotheses)" ((1987/1944), p. 31).

Consider the proposition that the logical consequence relation is reflexive (i.e., for any sentence X, X is a logical consequence of X). It is hard to grant that my indubitable knowledge that this proposition is true is not *a priori* for it seems unlikely that accumulated empirical experience accounts for my certainty of it. Furthermore, I can't even imagine the empirical circumstances that would disconfirm it. However, different examples may do better in motivating the thesis that logic is empirical. For example, *Kelly is at home and either playing her trumpet or reading a book* seems to entail that *Kelly is at home and playing her trumpet or at home and reading a book*. I can't imagine a time when the first sentence would be true and the second false. However, what holds universally for Kelly may not hold universally for quantum phenomena. It has been proposed that in order to make quantum mechanics consistent with the rest of physics, we reject that *quantum state A holds and either state B or C holds* entails *quantum states A and B hold or A and C hold* (Putnam (1969)). The issues here are substantive. We save further discussion of the *a prioricity* criterion for later in Chapter 4, and assume the criterion in the remainder of this chapter.

Let's summarize and tie things together. We began by asking, for a given language, what conditions must be met in order for a sentence X to be a logical consequence of a class K of sentences? Tarski thinks that an adequate response must reflect the common concept of logical consequence. By the lights of this concept, in order to fix what follows logically from what in a language, we must select a class of constants that determines an *a priori*, formal consequence relation which reflects the modal element in the common concept of logical consequence. Such constants are called *logical constants*, and we say that the sentential form of a sentence (i.e., its *logical* form) is a function of the logical constants that occur in the sentence and the pattern of the remaining expressions. As was illustrated above, the notion of formality does not *presuppose* a criterion of logical constancy. A consequence relation based on any division between constants and terms replaced with variables will automatically be formal with respect to that division of expressions.

Logical Constants

Tarski's characterization of the common concept of logical consequence in his (1936) paper rules out certain terms as logical. As illustrated above, treating as logical constants 'not' and predicates whose extensions are known only *a posteriori* such as "US President", and "Female" can result in a formal consequence relation that is not logical, because it fails the *a prioricity* criterion. Tarski does not explain in his (1936) paper what exactly counts as a logical constant. Without such an explanation, the class of logical constants is not completely determined, and this hampers our ability to clarify the rationale for a selection of terms to serve as logical. To elaborate, consider Tarski's remark on the boundary between logical and non-terms.

> Underlying this characterization of logical consequence is the division of all terms of the language discussed into logical and extra-logical. This division is not quite arbitrary. If, for example, we were to include among the extra-logical signs the implication sign, or the universal quantifier, then our definition of the concept of consequence would lead to results which obviously contradict ordinary usage. On the other had, no objective grounds are known to me which permit us to draw a sharp boundary between the two groups of terms. It seems to be possible to include among logical terms some which are usually regarded by logicians as extra-logical without running into consequences which stands in sharp contrast to ordinary usage. (1936, p. 419)

Tarski seems right to think that the logical consequence relation turns on the work that the logical terminology does in the relevant sentences. It seems odd to say that *Kelly is happy* does not logically follow from *All are happy* because the second is true and the first false when *All* is replaced with *Few*. However, by Tarski's version of the common concept of logical consequence there is no reason not to treat, say, *taller than* as a logical constant along with *not* and, therefore, no reason not to take *Kelly is not taller than Paige* as following logically from *Paige is taller than Kelly*. Also, it seems plausible to say that I know *a priori* that there is no possible interpretation of *Kelly* and *mortal* according to which *it is necessary that Kelly is mortal* is true and *Kelly is mortal* is false. This makes *Kelly is mortal* a logical consequence of *it is necessary that Kelly is mortal*. Given that *taller than* and *it is necessary that*, along with other terms, were not generally regarded as logical terms by logicians of the early-Tarski's day, the fact that they seem to be logical terms by the common concept of logical consequence, as observed by Tarski, highlights the question of what it takes to be a logical term. Tarski says that future research will either justify the traditional boundary between the logical and the non-logical or conclude that there is no such boundary and the concept of logical consequence is a relative concept whose extension is always relative to some selection of terms as logical (p.420).

In other work, Tarski is more informative about logical constancy. Consider the following two passages taken from the English translation of his introductory book in German on logic and general deductive methodology.

> The constants with which we have to deal in every scientific theory may be divided into two large groups. The first group consists of terms which are specific for a given theory. In the case of arithmetic, for instance, they are terms denoting either individual numbers or whole classes of numbers, relations between numbers, operations on numbers, … On the other hand, there are terms of a much more general character occurring in most of the statements of arithmetic, terms which are met constantly both in considerations of everyday life and in every possible field of science, and which represent an indispensible means for conveying human thoughts and for carrying out inferences in any field whatsoever; such words as "not", "and", "or", "is', "every", "some" and many others belong here. There is a special discipline, namely LOGIC, considered as the basis for all other sciences, whose concern is to establish the precise meaning of such terms and lay down the most general laws in which these terms are involved. ((1941) p.18)

> Among terms of a logical character there is a small distinguished group, consisting of such words as "*not*", "*and*", "*or*", "*if…, then…*". All these words are well-known

> from everyday language, and serve to build up compound sentences from simpler ones. In grammar, they are counted among the so-called sentential conjunctions. If only for this reason, the presence of these terms does not represent a specific property of any particular science. (p. 19)

According to Tarski, an essential characteristic of logical terms is their topic-neutrality, which reflects their indispensability as a "means for conveying human thoughts and for carrying out inferences in any field whatsoever." This commonly held view provides a *prima facie* reason for ruling out terms that are specific to a branch of inquiry (e.g., mathematical, physical, etc.) as logical constants. Logicians have traditionally regarded as logical constants the copula *is*, sentential connectives such as *and*, *not*, *or*, *if…then*, as well as the quantifiers *all* and *some*, and the identity predicate =. These terms are not only used in everyday life and in all fields of science, but it is hard to imagine how a thought could even be formulated without any of them. Admittedly, topic-neutrality is not a fine-grained criterion of logical constancy. It seems to rule out 'taller than', it is not obvious that 'it is necessary that' is ruled out.

A rationale for Tarski's view that the concern of logic is "to establish the precise meaning of [logical constants] and lay down the most general laws in which these terms are involved" may be put as follows. Since logical consequence is a type of formal consequence, whether or not the logical consequence relation obtains between a sentence X and a class K of sentences depends on the logical forms of X and the K-sentences. Recall that the logical form of a sentence is a function of the logical constants that occur in the sentence and the pattern of the remaining variable expressions that replace the non-logical expressions. Therefore, the logical consequence relation for a language L is determined, in part, by the meanings of L's logical constants, which are determinants of the logical forms of L-sentences. This makes it the case that logical necessity is grounded on the meanings of no expressions other than logical constants. Obviously, a language's logical consequence relation is logic's business, and so a concern of logic is to precisely understand aspects of the meanings of logical constants relevant to the determination of what follows from what. One question that arises is how, exactly, do the meanings of logical constants determine the logical consequence relation? In the next two sections, two distinct approaches to answering this question are sketched. Each approach leads to a distinctive understanding of logical necessity, and a unique way of characterizing logical consequence.

The truth-conditional properties of logical constants

Ignorant of US politics, I couldn't determine the truth of *Kelly is not US President* solely on the basis of *Kelly is a female*. However, behind such a veil of ignorance I would be able to tell that *Kelly is not US President* is true if *Kelly is female and Kelly is not US President* is true. How? Short answer: based on my linguistic competence; longer answer: based on my understanding of the contribution of *and* to the determination of the truth conditions of a sentence of the form *P and not-Q*. The truth conditions of a sentence are the conditions under which the sentence is true. For any sentences P and Q, I know that *P and not-Q* is true just in case P is true and *not-Q* is true. So, I know, *a priori*, if *P and not-Q* is true, then *not-Q* is true.

Taking *not* and *and* to be the only logical constants in (9) *Kelly is not both at home and at work*, (10) *Kelly is at home*, and (11) *Kelly is not at work*, we formalize the sentences as follows, letting the variable *k* replace *Kelly*, and predicate variables *H* and *W* replace *is at home*, and *is at work*, respectively.

9'. not-(Hk and Wk)
10'. Hk
11'. not-Wk

Read *Hk* as *k is an H* and *Wk* as *k is a W*. There is no interpretation of *k*, *H*, and *W* according to which (9') and (10') are true and (11') is false. The reason why turns on the meanings of *and* and *not*, which are knowable *a priori*. To elaborate, the linguistic (as opposed to contextual) meanings of logical constants determine their truth-conditional properties, i.e., determine what they contribute to the truth conditions of declarative sentences in which they occur (in extensional contexts). The truth-conditional property of a logical constant is a semantic property that seems to be determined *a priori*. For example, the traditional understanding of the truth-conditional properties of 'not' and 'and' may be described as follows: For any sentences P and Q, the truth-value of *not-P* is the reverse of that of P, and *P and Q* is true if P is true and Q is true and false if P is false or Q is false. Arguably, empirical experience does not seem relevant to the justification of this view of the truth-conditional properties of 'not' and 'and'.

Suppose (9') and (10') are true on some interpretation of the variable terms. Then the truth-conditional property of *not* in (9') makes it the case that *Hk and Wk* is false, which, in accordance with the truth-conditional property of *and* requires that *Hk* is false or *Wk* is false. Given (10'), it must be that *Wk* is false, i.e., *not-Wk* is true. So, there can't be an interpretation of the variable expressions according to which (9') and (10') are true and (11') is false, and, as the above reasoning illustrates, this is due exclusively to the truth-conditional properties of *not* and *and*. The reason that an interpretation of the variable expressions according to which (9') and (10') are true and (11') is false is impossible is that the supposition otherwise is inconsistent with the meanings of *not* and *and*, as they determine the truth-conditional properties of these expressions. Compare: the supposition that there is an interpretation of *k* according to which *k is a female* is true and *k is not US President* is false does not seem to violate the truth-conditional properties of the constant terms. If we identify the truth-conditional properties of the predicates with their extensions in all possible worlds, then the supposition that there is a female U.S. President does not violate the meanings of *female* and *US President* for surely it is possible that there be a female US President. But, supposing that (9') and (10') could be true with (11') false on some interpretation of *k*, *H*, and *W*, violates the truth-conditional properties of *and* or *not*.

It is logically necessary that if (9) and (10) are true, then (11) is true because, given the pattern of the non-logical terms, the supposition that (9) and (10) are true and (11) is false is incompatible with the truth-conditional properties of *and* and *not*. Since the truth-conditional property of a logical constant is a semantic property determined by its meaning, we may say that the possibility of (9) and (10) being true and (11) false is ruled out by the meanings of the component logical terms. A preliminary notion of logical necessity on the truth-conditional approach to the meanings of the logical constants may be characterized succinctly as follows. It is logically necessary that if a class K of sentences is true, then a logical consequence X of K is true, because the possibility of all true K-sentences and a false X is ruled out by the truth-conditional properties of occurrent logical terms and the pattern of non-logical terms.

We began with Tarski's characterization of what he calls the common concept of logical consequence as (i) necessary, (ii) formal, and (iii) knowable *a priori*. Our rendition of (i)-(iii), yields,

> A sentence X is a logical consequence in a language L of a class K of sentences iff there is no possible interpretation of the non-logical terms of L that makes all of the sentences in K true and X false.

Let's call this precisification of the common concept of logical consequence, the *Tarskian precising definition of logical consequence*. It seems wrong to view this definition as pre-theoretical. For example, Tarski's account of formal consequence is theoretical, and we have employed its machinery to derive a notion of logical necessity. We have taken what Tarski calls the common concept of logical consequence to be the concept of logical consequence as it is employed by typical reasoners. Is it really true that the concept of logical consequence as typically employed is necessary and formal in the ways adumbrated above? Perhaps we should ignore Tarski's suggestions that his characterization of the common concept of logical consequence is intended to adhere to its employment by the typical reasoner, and regard it, as some have argued, (e.g., Hodges (2007) Jané (2006)), as Tarski's attempt to reflect the concept as employed by researchers in the foundations of mathematics. Putting this issue aside, it is clear that Tarski's precising definition reflects Tarski's idea that the use of the common concept to determine the logical consequence relation for a given language L presupposes a *prior* distinction between the logical and non-logical terms of L.

A possible interpretation of the non-logical terms of L according to which sentences are true or false is a reading of them according to which the sentences receive a truth-value in a situation that is not ruled out by the truth-conditional properties of the logical constants. The philosophical locus of the technical, theoretical development of 'possible interpretation' in terms of models is Tarski (1936).

Tarski defines the *satisfaction* of a sentential form by an object or sequence of objects (e.g., John and Peter may satisfy 'x and y are brothers'; the sequence 2, 3, 5 satisfies 'x + y = z', the sequence 3, 5, 2 does not). Then Tarski defines a *model* or a *realization* of a class L of sentences as a sequence of objects satisfying all the logical forms derived from the sentences of L by replacing all extra-logical constants which occur in the sentences belonging to L by corresponding variables, like constants being replaced by like variables and unlike by unlike (pp. 416-17). A model of a sentence is a model of the class consisting of just the sentence. Tarski's theoretical definition of logical consequence is as follows.

> The sentence X follows logically from the sentences of the class K iff every model of the class K is also a model of the sentence X. (p. 417)

Models have become standard tools for characterizing the logical consequence relation, and the characterization of logical consequence in terms of models is called the Tarskian or model-theoretic characterization of logical consequence. We say that X is a model-theoretic consequence of K iff all models of K are models of X. This relation may be represented as K $\models$ X.

Immediately after giving his model-theoretic definition of logical consequence, Tarski asserts that, "[i]t seems to me that everyone who understands the content of the above definition must admit that it agrees quite well with common usage" (p. 418). Tarski assertion, which he never defends anywhere in his writings, raises the issue of how a theoretical definition of logical consequence such as the model-theoretic definition is to be justified. This issue will be discussed in Chapter 4, after we have presented in detail a model-theoretic definition of logical consequence. Here we merely note the following. There are three concepts of logical consequence at play in Tarski's (1936) article: (i) a common concept, (ii) a precise concept, and (iii) a model-theoretic concept. Tarski's assertion, understood as claiming that (iii) adequately represents (i), raises the issue of the roles (i) and (ii) play in justifying (iii) as an adequate representation of logical consequence. For example, we might think that (iii) is inadequate in one of two ways: (iii) does not represent (ii), or it does but (iii) doesn't represent (i) because (ii) is an inadequate representation of (i).

The inferential properties of logical constants

We now turn to a second approach to the meanings of logical constants. Instead of viewing the meanings of logical constants as determinants of their truth-conditional properties as is done on the model-theoretic approach, on the second approach we view the meanings of logical constants as determinants of their inferential properties conceived of in terms of principles of inference, i.e. principles justifying steps in deductions. We begin with a remark made by Aristotle. In his study of logical consequence, Aristotle comments that,

> a syllogism is discourse in which, certain things being stated, something other than what is stated follows of necessity from their being so. I mean by the last phrase that

they produce the consequence, and by this, that no further term is required from without in order to make the consequence necessary. (*Prior Analytics* 24b)

Adapting this to our X and class K, we may say that X is a logical consequence of K when the sentences of K are sufficient *to produce* X. Since the K-sentences are sufficient to produce X, it is necessarily the case that if all the former are true, then X is true as well. How are we to think of a sentence being produced by others? One way of developing this is to appeal to a notion of an actual or possible deduction. X is a deductive consequence of K iff there is a deduction of X from K. In such a case, we say that X may be correctly inferred from K, or that it would be correct to conclude X from K, if the K–sentences are true. A deduction is associated with an ordered pair <K, X> (see Chapter 3 for an explanation of ordered pairs); the class K of sentences is the basis of the deduction, and X is the conclusion. A deduction from K to X is a finite sequence S of sentences ending with X such that each sentence in S (i.e., each intermediate conclusion) is derived from a sentence (or more) in K or from previous sentences in S in accordance with a correct principle of inference, i.e., a correct production rule. For example, intuitively, the following inference seems correct.

9. It is not true that Kelly is both at home and at work
10. <u>Kelly is at home</u>
∴ 11. It is not true that Kelly is at work

The symbol '∴' means 'therefore'. The class K of sentences above the line is the basis of the inference, and the sentence X below is the conclusion. We represent their logical forms as follows.

9'. not-(Hk and Wk)
10'. <u>Hk</u>
∴ 11'. not-Wk

Consider the following deduction of (11') from (10') and (9').

Deduction: Assume that (12') *Wk*. Then from (10') and (12') we may deduce that (13') *Hk and Wk*. (13') contradicts (9') and so (12'), our initial assumption, must be false. We have deduced *not-Wk* from *not-(Hk and Wk)* and *Hk.*

Since the deduction of *not-Wk* from *not-(Hk and Wk)* and *Hk* did not depend on the interpretation of *k*, *W*, and *H*, the deductive relation is formal. Furthermore, my knowledge of this is *a priori* because my non-empirical knowledge of the underlying principles of inference in the above deduction. For example, letting P and Q be any sentences, we know *a priori* that *P and Q* may be inferred from the class K={P, Q} of basis sentences. This inferential principle derived from the meaning of *and* grounds the move from (10') and (12') to (13'). Also, the deduction appeals to the principle that if we deduce a contradiction from an assumption P, then we may infer that the assumption is false, i.e., we may infer that not-P. The correctness of this principle, derived from the meaning of *not*, seems to be an *a priori* matter. Here we view the meaning of a logical constant as fixing what can be inferred from a sentence in which it is the dominant constant, and as determining the types of premises from which such a sentence can be immediately concluded when the premises are true. Note that the above deduction illustrates that (11) follows of necessity from (9) and (10) because the latter produce the former. It is logically necessary that if (9) and (10) are true, then (11) is true because there is deduction from (9) and (10) to (11). Let's look at another example of a deduction.

1. Some children are both lawyers and peacemakers
∴ 2. Some children are peacemakers

The logical forms are of (1) and (2) are represented as follows.

1'. Some S are both M and P
∴ 2'. Some S are P

Again, intuitively, (2') is deducible from (1').

> Deduction: The basis tells us that at least one *S*—let's call this *S* '*a*'—is an *M* and a *P*. Clearly, *a is P* may be deduced from *a is an M and a P*. Since we've assumed that *a is an S*, what we derive with respect to *a* we derive with respect to some *S*. So our derivation of *a is an P* is a derivation of *Some S is an P*, which is our desired conclusion.

Since the deduction is formal, we have shown not merely that (2) can be correctly inferred from (1), but we have shown that for any interpretation of *S*, *M*, and *P* it is correct to infer (2') from (1').

Typically, deductions leave out steps (perhaps because they are too obvious), and they usually do not justify each and every step made in moving towards the conclusion (again, obviousness begets brevity). The notion of a deduction is made precise by describing a mechanism for constructing deductions that are both transparent and rigorous (each step is explicitly justified and no steps are omitted). This mechanism is a deductive system (also known as a formal system or as a formal proof calculus). A deductive system D is a collection of rules that govern which sequences of sentences, associated with a given <K, X>, are allowed and which are not. Such a sequence is called a proof in D (or, equivalently, a deduction in D) of X from K. The rules must be such that whether or not a given sequence associated with <K, X> qualifies as a proof in D of X from K is decidable purely by inspection and calculation. That is, the rules provide a purely mechanical procedure for deciding whether a given object is a proof in D of X from K.

We say that a deductive system D is correct when for any K and X, proofs in D of X from K corresponds to deductions. For example, intuitively, there are no correct principles of inference according to which it is correct to conclude

Some animals are mammals and reptiles

on the basis of the following two sentences.

Some animals are mammals
Some animals are reptiles

Hence, a proof in a deductive system of the former sentence from the latter two is evidence that the deductive system is incorrect. The point here is that a proof in D may fail to represent a deduction if D is incorrect.

A rich variety of deductive systems have been developed for registering deductions. Each system has its advantages and disadvantages, which are assessed in the context of the more specific tasks the deductive system is designed to accomplish. Historically, the general purpose of the construction of deductive systems was to reduce reasoning to precise mechanical rules (Hodges 1983, p. 26). Some view a deductive system defined for a language L as a mathematical model of actual or possible chains of correct reasoning in L.

If there is a proof of X from K in D, then we may say that X is a deductive consequence in D of K, which is sometimes expressed as $K \vdash_D X$. Relative to a correct deductive system D, we characterize logical consequence in terms of deductive consequence as follows.

X is a logical consequence of K iff X is a deductive consequence in D of K, i.e. there is an actual or possible proof in D of X from K.

This is sometimes called the proof-theoretic characterization of logical consequence. Tarski calls it the formalized concept of consequence in his (1936). He favors a model-theoretic approach to logical consequence. We shall consider Tarski's reasoning in Chapter 5, where we give a deductive system for a sample language and consider why logical consequence might not be equivalent with deducibility in a deductive system.

The Model-Theoretic and Deductive-Theoretic Approaches to Logic

Let's step back and summarize where we are at this point in our discussion of the logical consequence relation. We began with Tarski's observations of the common or ordinary concept of logical consequence. According to Tarski, if X is a logical consequence of a class of sentences, K, then, in virtue of the logical forms of the sentences involved, necessarily if all of the members of K are true, then X is true, and furthermore, we know this *a priori*. The formality criterion makes the meanings of the logical constants and the pattern of the remaining non-logical terms the essential determinants of the logical consequence relation. We highlighted two different approaches to the meanings of logical constants: (1) in terms of the constant's truth-conditional properties, and (2) in terms of their inferential properties. Each yields a distinct conception of the notion of necessity inherent in the common concept of logical consequence, and leads to the following two characterizations of logical consequence.

1. X is a logical consequence of K iff there is no possible interpretation of the non-logical terminology of the language according to which all the sentence in K are true and X is false.

2. X is a logical consequence of K iff X is deducible from K.

We make the notions of *possible interpretation* in (1) and *deducibility* in (2) precise by appealing to the technical notions of *model* and *deductive system*. This leads to the following theoretical characterizations of logical consequence.

The *model-theoretic characterization of logical consequence*: X is a logical consequence of K iff all models of K are models of X.

The *deductive- (or proof-) theoretic characterization of logical consequence*: X is a logical consequence of K iff there is a deduction in a deductive system of X from K.

We said in Chapter 1 that the primary aim of logic is to tell us what follows logically from what. These two characterizations of logical consequence lead to two different orientations or conceptions of logic (see Tharp (1975), p.5).

> *Model-theoretic approach*: Logic is a theory of possible interpretations. For a given language the class of situations that can—logically—be described by that language.
> *Deductive-theoretic approach*: Logic is a theory of formal deductive inference.

Following Shapiro ((1991) p.3) define a logic to be a language L plus either a model-theoretic or a proof-theoretic account of logical consequence. A language with both characterizations is a *full logic* just in case both characterizations coincide. In Chapter 3, we define a language, and then characterize the logical consequence relation in this language model-theoretically and proof theoretically in Chapters 4 and 5, respectively.

Chapter 3
Set-Theoretic and Linguistic Preliminaries

The model-theoretic and deductive-theoretic characterizations of logical consequence make essential use of concepts and terminology from set theory. In this chapter, we introduce the informal notion of a set, and explain the terminology from set theory that shall be used in later chapters. The reader may skip these explanations now, and if necessary refer to them later on as the terminology and concepts are used. Also, in this chapter we define a simple language M for which we will later characterize logical consequence, and introduce some important syntactic and semantic concepts.

Set-Theoretic Preliminaries

A set is an aggregate or collection of entities. Which aggregates form sets is an issue that we shall touch on later. We employ the *epsilon* notation for set membership: '$x \in y$' is read 'x is a member of y' or 'x is an element of y'. '$x \notin y$' means 'x is not a member of y'. A set may be specified by listing its members and enclosing the list in curly brackets. Given any entities $x_1, \ldots, x_n$, $\{x_1, \ldots, x_n\}$ is the set that contains just $x_1, \ldots, x_n$. For example, {2, 3} is the set whose sole members are 2 and 3. The set {{2, 3}} has the set {2, 3} as its only member. Sets S and S′ are identical iff S and S′ have exactly the same members. Thus, {Socrates, Plato} and {Socrates, Socrates, Plato} are identical with {Plato, Socrates}. A set S is a subset of a set S′ iff every member of S is a member of S′. Both {Socrates} and {Socrates, Plato} are subsets of {Socrates, Plato}.

By $<x_1, \ldots, x_n>$, we mean the ordered *n*-tuple of $x_1, \ldots, x_n$. The essential property of ordered *n*-tuples is that $<x_1, \ldots, x_n> = <y_1, \ldots, y_n>$ iff $x_1 = y_1$, $x_2 = y_2$, … , $x_n = y_n$. Thus, <Socrates, Plato> ≠ <Plato, Socrates>. Ordered 2-tuples are called ordered pairs. If S is a set and *n* an integer greater than 1, then S^n is the set of all ordered n-tuples $<x_1, \ldots, x_n>$ of elements $x_1, \ldots, x_n$ of S. The Cartesian product of sets S and S′ (denoted by $S \times S'$) is the set of all ordered pairs $<x, y>$ such that $x \in S$ and $y \in S'$. Letting *n* be a positive integer, the set S^n is the Cartesian product of S with itself *n* times. Set S^1 is set S, and S^2 is the set of all ordered pairs of members of S, S^3 is the set of all ordered triples of members of S, etc.

The extension of a relation is the set of entities the relation is true of . An n-place relation on a set S is a subset of S^n. A 1-place relation R on a set S (called a property on S) is the set of *x*s such that $x \in S$ and $x\in R$ ('$x\in R$' means *x is a member of the extension of R*). For example, *being human* on the set of mammals is the set of x*s* such that *x* is a mammal and *x* is a human. A 2-place relation R (called a *binary* relation) on a set S is the set of all ordered pairs $<x,y>$ such that $x \in S$ and $y \in S$, and $<x,y>\in R$ (i.e., the ordered pair $<x,y>$ is a member of the extension of R, and so x is R-related to y in that order). For example, *being the biological mother of* on the set of human beings is the set of ordered pairs $<x,y>$ such that *x* and *y* are human beings and the biological mother of x is *y*. Note that the extension of a binary relation on a set S is a subset of S^2.

We may represent the extension of a property P as, $\{x\in S \mid P(x)\}$. This is the set S of *x*s that P is true of (e.g., $\{x\in S|$ x is a prime number less than four$\} = \{2, 3\}$). The set $\{x\in S| x \neq x\}$ is the empty set, which is denoted by the symbol '$\varnothing$'. As indicated above, the extension of a relation R on a set S may be an ordered *n*-tuple. The set $\{<x_1, \ldots, x_n> \in S^n| R(x_1, \ldots, x_n)\}$ denotes the extension of R. For example, $\{<x, y>\in N^2| x < y\}$ is the extension of the $<$-relation on the set of natural numbers. If a relation R has an extension (i.e., if the collection of entities that it is true of is a set), then we may say that R determines a set.

The *domain* of a binary relation R on a set S is the set of all x such that $<x, y> \in R$ for some y. The *range* of R is the set of all *y* such that $<x, y> \in R$ for some x. A function *f* is a binary relation such that if $<x, y>$ and $<x,z>$ are members of its extension, then $y = z$. Therefore, for any member of *f* 's domain, there is a unique *y* such that $<x,y> \in f$. The expression '$f(x)$' (*x* is the argument of *f*) refers to this unique element *y*. Consider a relation *f* whose domain is the set of positive integers, its range the set of even integers, and is defined by $f(x) =2x$. The relation *f* is a function; for each integer *x* is assigned one unique number that is the product *x* and 2. If *n*-tuples are the elements of the domain of a function *f*, then *f* is a function of *n*-arguments. The relation *f* whose domain is the set of 2-tuples of integers, whose range is the set of integers, and is defined by $f(<x, y>) = x + y$, is a function of 2-arguments. The expression $f(<x_1, \ldots, x_n>)$ may be written as $f(x_1, \ldots, x_n)$, which is the value of *f* for arguments $x_1, \ldots, x_n$.

Linguistic Preliminaries

We now characterize a language M by sketching what strings qualify as well-formed formulas (also called *formulas* or *wffs*) in M, and then we define sentences from wffs. Next we give an account of truth in M, i.e. we describe the conditions in which M-sentences are true.

Syntax of M

Building blocks of formulas

Terms

Individual names—'beth', 'kelly', 'matt', 'paige', 'shannon', 'evan', and 'w_1', 'w_2', 'w_3', etc.

Variables—'x', 'y', 'z', 'x_1', 'y_1', 'z_1', etc.

Predicates

1-place predicates—'Female', 'Male'

2-place predicates—'Parent', 'Brother', 'Sister', 'Married', 'Older-Than', 'Admires', '='.

Blueprints of well-formed formulas (wffs)

(1) An atomic wff is any of the above *n*-place predicates followed by *n* terms which are enclosed in parentheses and separated by commas.

Sentential Connectives (~, &, ∨, →)

(2) If α is a wff, so is ~ α.
(3) If α and β are wffs, so is (α & β).
(4) If α and β are wffs, so is ($\alpha \vee \beta$).
(5) If α and β are wffs, so is ($\alpha \rightarrow \beta$).

Quantifiers (∃,∀)

(6) If Ψ is a wff and *v* is a variable, then $\exists v\Psi$ is a wff.
(7) If Ψ is a wff and *v* is a variable, then $\forall v\Psi$ is a wff.

Finally, no string of symbols is a formula of M unless the string can be derived from the formation rules (1)-(7). It should be noted that the Greek letters used in stating (2)-(7) are not symbols of M (more on this later in the chapter).

It will prove convenient to have available in M an infinite number of individual names as well as variables. The strings 'Parent(beth, paige)' and 'Male(x)' are examples of atomic wffs. We allow the identity symbol in an atomic formula to occur in between two terms, e.g., instead of '=(evan, evan)' we allow (evan = evan). The symbols '~', '&', '∨', and '→' correspond to the English words 'not', 'and', 'or' and 'if…then', respectively. '∃' is our symbol for an existential quantifier and '∀' is the universal quantifier. ∃vΨ and ∀vΨ correspond to *for some v, Ψ* and *for all v, Ψ*, respectively. For every quantifier, its scope is the smallest part of the wff in which it is contained that is itself a wff. An occurrence of a variable *v* is a bound occurrence iff it is in the scope of some quantifier '∃' or '∀', and is free otherwise. For example, the occurrence of *x* is free in 'Male(x)' and in '∃yMarried(y,x)'. The occurrence of *y* in the second formula is bound because it is in the scope of the existential quantifier. A wff with at least one free variable is an open wff, and a closed wff is one with no free variables. A sentence is a closed wff. For example, 'Female(kelly)' and '∃y∃xMarried(y,x)' are sentences but 'OlderThan(kelly,y)' and '(∃xMale(x) & Female(z))' are not.

Recall that a string of M's symbols is a wff of M iff it is derivable by the formation rules (1)-(7). Whether such a string is so derivable is independent of any interpretation of the string. For example, that '~ Parent(beth, paige)' is derivable from (1) and then (2) is independent of any interpretation of its sentential connective and terms. The derivation of a wff from (1)-(7) is a unique recipe for putting it together. Each wff α can be put together only in one way. Either α is an atomic formula or there is one unique clause among (2)-(7) that was last applied to produce α (no formula produced by (2)-(7) is atomic). We say that the last clause applied in deriving a wff produces the wff. If either (2), (3), (4), or (5) was last applied, then α is a negation, conjunction, disjunction, or conditional formula, respectively. If (6) or (7) was last applied, then α is either an existential or universal quantification. No wff can be produced by two different clauses.

Language M counts as a formal language, because M is a language that can be completely defined without any reference to its interpretation (Hunter 1971, p. 4) even though it is natural to see it as an instrument for communicating facts about a family. More specifically, a formal language can be identified with the set of its wffs, and is specified by giving its alphabet of symbols (M's alphabet is its terms, sentential connectives, quantifiers, comma, and open and closed parentheses), and its formation rules (e.g., M's (1)-(7)), which determine which strings constructible from the alphabet are wffs. Since both the alphabet and formation rules can be completely specified without any reference to interpretation, a formal language can be completely characterized syntactically. A formal language is sometimes referred to as an artificial language in order to distinguish it from a natural language, i.e., one spoken and written as the native language of human beings.

Before moving to semantics, we note that the expressions of M's alphabet are orthographic or phonographic types, consisting respectively of sequences of mark or phone—simple vowel or consonant sound—types. Their utterances/inscriptions are their tokens. Two utterances/inscriptions are occurrences of an alphabet expression iff they are tokens of the same type of expression. There can be spoken and written variants of an expression (e.g., variations in pronunciation and variations in sign/mark design), and so we need to conceive of an expression from one language as a set of conventionally related types from different spoken and written variations of that language. Furthermore, since expressions can be (and are) in the alphabets of different languages, expressions must be taken to be something like a set of conventionally related types from different first-order languages. Borrowing from (Alward, (2005)), we illustrate these two observations as follows.

Let *a* be an element of a first-order language L's alphabet. Lw_1, Lw_2, …, Lw_n are written variants of L's alphabet, and Ls_1, Ls_2, …,Ls_n are spoken variants of L's alphabet. The ordered pairs $<a_{w_i}, Lw_i>$ and $<a_{s_i}, Ls_i>$ designate types of inscriptions and sounds, respectively, relative to written and spoken variants of language L. We represent *a* relative to L (i.e., $a\in L$) as follows:

$$a\in L= \{<a_{w_1}, Lw_1>, \ldots <a_{w_n}, Lw_n>, <a_{s_1}, Ls_1>, \ldots <a_{s_1}, Ls_1>\}$$

$a\in L$ is a collection of orthographic and phonographic types relativized to L. Any token of an element of $a\in L$ is a token of $a\in L$. The alphabet expression *a* may be defined extensionally as the collection of such sets.

$$a=\{ a\in L_1, a\in L_2, \ldots, a\in L_n \}$$

Two expressions are identical iff the collections that are definitive of them have the same members. A token of *a* is a token of any $a \in L_i$, and so a token of *a* is an inscription or an utterance. Two inscriptions/utterances are occurrences of the same expression *a* just in case they are tokens of *a*. Since an expression is an orthographic or phonographic type, being composed of respectively sequences of particular mark or phone types is essential to it. In Chapter 4, we shall appeal to this understanding of the essence of an expression to explain why the meaning assigned to an expression is an accidental feature of it. A more comprehensive account of expressions, would have to spell out, among other things, the properties tokens must possess in order to be of the same type a_{w_i} or type a_{s_i}, and the properties the latter must possess in order to be variants of a.

A language's wffs, composed of expressions from the alphabet, are also orthographic or phonographic types. Two utterances/inscriptions are occurrences of a wff iff they are tokens of the same type of wff. Since there can be spoken and written variants of a wff and one wff can be in different languages, the nature of a wff may be pictured along the lines of the above portrayal of alphabet-expressions.

Semantics for M

A semantics for a formal language provides the necessary and sufficient truth conditions for each sentence of the language. The truth conditions for a sentence are the circumstances under which the sentence is true. We now provide a semantics for our language M. This is done in two steps. First, we provide an interpretation of M by specifying a domain of discourse, i.e., the chunk of the world that M is about, and interpreting M's predicates and names in terms of the elements composing the domain. Then we state the conditions under which each type of M-sentence is true. To each of the above formation rules (1-7) there corresponds a semantic rule that stipulates the conditions under which the sentence constructed using the formation rule is true. We assume the principle of bivalence which states that there are just two truth-values, true and false, and all sentences are one or the other but not both. So, 'not true' and 'false' may be used interchangeably. In effect, the interpretation of M determines a truth-value for each and every sentence of M.

Domain D—The set of McKeon's: Matt, Beth, Shannon, Kelly, Paige, and Evan.

Here are the referents and extensions of the names and predicates of M.

> **Terms**: 'matt' refers to Matt, 'beth' refers to Beth, 'shannon' refers to Shannon, etc…

A relation is identified with its extension, and the meaning of a predicate is the extension of the relation on D that it designates. Thus the interpretation of a one-place predicate is a set of elements from D, and the interpretation of a two-place predicate is a set of ordered pairs of elements from D.

> The extension of 'Male' is {Matt, Evan}
>
> The extension of 'Female' is {Beth, Shannon, Kelly, Paige}
>
> The extension of 'Parent' is {<Matt, Shannon>, <Matt, Kelly>, <Matt, Paige>, <Matt, Evan>, <Beth, Shannon>, <Beth, Kelly>, <Beth, Paige>, <Beth, Evan>}
>
> The extension of 'Married' is {<Matt, Beth>, <Beth, Matt>}
>
> The extension of 'Sister' is {<Shannon, Kelly>, <Kelly, Shannon>, <Shannon, Paige>, <Paige, Shannon>, <Kelly, Paige>, <Paige, Kelly>, <Shannon, Evan>, <Kelly, Evan>, <Paige, Evan>}
>
> The extension of 'Brother' is {<Evan, Shannon>, <Evan, Kelly>, <Evan, Paige>}
>
> The extension of 'OlderThan' is {<Beth, Matt>, <Beth, Shannon>, <Beth, Kelly>, <Beth, Paige>, <Beth, Evan>, <Matt, Shannon>, <Matt, Kelly>, <Matt, Paige>, <Matt, Evan>, <Shannon, Kelly>, <Shannon, Paige>, <Shannon, Evan>, <Kelly, Paige>, <Kelly, Evan>, <Paige, Evan>}
>
> The extension of 'Admires' is {<Matt, Beth>, <Shannon, Matt>, <Shannon, Beth>, <Kelly, Beth>, <Kelly, Matt>, <Kelly, Shannon>, <Paige, Beth>, <Paige, Matt>, <Paige, Shannon>, <Paige, Kelly>, <Evan, Beth>, <Evan, Matt>, <Evan, Shannon>, <Evan, Kelly>, <Evan, Paige>}
>
> The extension of '=' is {<Matt, Matt>, <Beth, Beth>, <Shannon, Shannon>, <Kelly, Kelly>, <Paige, Paige>, <Evan, Evan>}

(I) An atomic sentence with a one-place predicate is true iff the referent of the term is a member of the extension of the predicate, and an atomic sentence with a two-place predicate is true iff the ordered pair formed from the referents of the terms in order is a member of the extension of the predicate.

The atomic sentence ‘Female(kelly)’ is true because, as indicated above, the referent of ‘kelly’ is in the extension of the property designated by ‘Female’. The atomic sentence ‘Married(shannon, kelly)’ is false because the ordered pair <Shannon, Kelly> is not in the extension of the relation designated by ‘Married’.

Let α and β be any M-sentences.

(II) ~α is true iff α is false.

(III) (α & β) is true when both α and β are true; otherwise (α & β) is false.

(IV) (α ∨ β) is true when least one of α and β is true; otherwise (α∨ β) is false.

(V) (α→β) is true iff α is false or β is true. So, (α→ β) is false just in case α is true and β is false.

(I)–(V) depict the truth-conditional properties of terms, predicates, and the sentential connectives. The symbols ‘~’ and ‘&’, thus interpreted, roughly correspond to ‘not’ and ‘and’ as ordinarily used. We call ~α and (α & β) negation and conjunction formulas, respectively. The formula (α∨ β) is called a disjunction and the truth-conditional property of ‘∨’ corresponds to that of inclusive-or. There are a variety of conditionals in English (e.g., causal, counterfactual, logical), each type having a distinct meaning. The conditional defined by (V) is called the material conditional. One way of following (V) is to see that the truth conditions for (α→β) are the same for ~(α & ~β).

By (II) ‘~Married(shannon, kelly)’ is true because, as noted above, ‘Married(shannon,kelly)’ is false. (II) also tells us that ‘~Female(kelly)’ is false since ‘Female(kelly)’ is true. According to (III), ‘(~Married(shannon,kelly) & Female(kelly))’ is true because ‘~Married(shannon, kelly)’ is true and ‘Female(kelly)’ is true. And ‘(Male(shannon) & Female(shannon))’ is false because ‘Male(shannon)’ is false. (IV) confirms that ‘(Female(kelly) ∨

Married(evan, evan))' is true because, even though 'Married(evan, evan)' is false, 'Female(kelly)' is true. From (V) we know that the sentence '(~(beth=beth) → Male(shannon))' is true because '~(beth=beth)' is false. If α is false then (α→β) is true regardless of whether or not β is true. The sentence '(Female(beth) → Male(shannon))' is false because 'Female(beth)' is true and 'Male(shannon)' is false.

Before describing the truth-conditional properties of the quantifiers we need to say something about the notion of satisfaction. We've defined truth only for the formulas of M that are sentences. So, the notions of truth and falsity are not applicable to non-sentences such as 'Male(x)' and '((x=x) → Female(x))' in which *x* occurs free. However, objects may satisfy wffs that are non-sentences. We introduce the notion of satisfaction with some examples. An object satisfies 'Male(x)' just in case that object is male. Matt satisfies 'Male(x)', Beth does not. This is the case because replacing 'x' in 'Male(x)' with 'matt' yields a truth while replacing the variable with 'beth' yields a falsehood. An object satisfies '((x=x) →Female(x))' iff it is either not identical with itself or is a female. Beth satisfies this wff (we get a truth when 'beth' is substituted for the variable in all of its occurrences), Matt does not (putting 'matt' in for 'x' wherever it occurs results in a falsehood). As a first approximation, we say that an object with a name, say 'a', satisfies a wff Ψ*v* in which at most *v* occurs free iff the sentence that results by replacing *v* in all of its occurrences with 'a' is true. 'Male(x)' is neither true nor false because it is not a sentence, but it is either satisfiable or not by a given object. Now we define the truth conditions for quantifications, utilizing the notion of satisfaction. The notion of satisfaction will be revisited later in Chapter 4 when we formalize the semantics for M and give the model-theoretic characterization of logical consequence.

Let Ψ be any formula of M in which at most v occurs free in Ψ.

(VI) ∃vΨ is true just in case there is at least one individual in the domain of quantification (e.g. at least one McKeon) that satisfies Ψ .

(VII) ∀vΨ is true just in case every individual in the domain of quantification (e.g. every McKeon) satisfies Ψ.

Here are some examples. '∃x(Male(x) & Married(x,beth))' is true because Matt satisfies '(Male(x) & Married(x,beth))'; replacing 'x' wherever it appears in the wff with 'matt' results in a true sentence. The sentence '∃xOlderThan(x,x)' is false because no McKeon satisfies 'OlderThan(x,x)', i.e., replacing 'x' in 'OlderThan(x,x)' with the name of a McKeon always yields a falsehood.

The universal quantification '∀x(OlderThan(x,paige) →Male(x))' is false for there is a McKeon who doesn't satisfy '(OlderThan(x,paige) → Male(x))'. For example, Shannon does not satisfy '(OlderThan(x,paige) → Male(x))', because Shannon satisfies 'OlderThan(x,paige)' but not 'Male(x)'. The sentence '∀x(x=x)' is true, because all McKeons satisfy 'x=x'; replacing 'x' with the name of any McKeon results in a true sentence.

Note that in the explanation of satisfaction we suppose that an object satisfies a wff only if the object is named. But we don't want to presuppose that all objects in the domain of discourse are named. For the purposes of an example, suppose that the McKeon's adopt a baby boy, but haven't named him yet. Then, '∃xBrother(x,evan)' is true because the adopted child satisfies 'Brother(x,evan)', even though we can't replace 'x' with the child's name to get a truth. To get around this is easy enough. We have added a list of names, 'w_1', 'w_2', 'w_3 ', etc. to M, and we may say that any unnamed object satisfies Ψ*v* iff the replacement of *v* with a previously unused w_i assigned as a name of this object results in a true sentence. In the above scenerio, '∃xBrother(x,evan)' is true, because treating 'w_1' as a temporary name of the child, 'Brother(w_1,evan)' is true.

We have characterized an interpreted formal language M by defining syntactically what qualifies as a sentence of M, and by giving the truth definition for M. In giving the truth condition for M we have characterized the truth-conditional properties of terms, connectives and quantifiers, and thereby characterized the conditions under which any M sentence is true. It is standard to distinguish between the language that is the object of study (the one mentioned) from the language in which the study is carried out (the one used). The former is called the object language and the later the metalanguage. Language M is our object language, and, as can be seen, the metalanguage is English, supplemented with symbols not in M's alphabet such as expressions from set theory, Greek letters (α, β Ψ), and the letter *v*. One motivation for distinguishing between object and metalanguages may be developed from Tarski's theory regarding the adequacy criteria for definitions of truth (Tarski (1933), Tarski (1944)).

For formal language L and its metalanguage L′, an instance of the schema

… is true in L iff _____

is a sentence of L′ such that '…' is replaced with a quotation-mark name of a sentence from L, and '____' is replaced with a translation in L′ of that sentence. Following Tarski, the schema may be called the T-schema. A definition of truth for L is materially adequate only if every instance of the T-schema is provable in L′ on the basis of the truth-definition and principles of logic. For example, the sentence,

> '∀x(OlderThan(x,paige) →Male(x))' is true in M iff every McKeon older than Paige is male

is an instance of the T-schema, and the material adequacy criterion requires that it be provable in M's metalanguage. This is provable since the above truth definition for M is materially adequate.

A definition of truth for a formal language must not only be materially adequate, but also formally correct, i.e., not entail a contradiction. Tarski believed that the procedure for obtaining a formally correct definition of truth for a formal language L requires that any expression functionally equivalent with 'is true in L' (i.e., the truth predicate for L) be excluded from L, and included, along with its definition, in L's metalanguage. If we let an object language L serve as its own metalanguage so that L contains names of its sentences, and 'is true in L' as a predicate, then, assuming elementary logic, the material adequacy condition for L's definition of truth entails a contradiction. For example, suppose that S_1 and S_2 are L-sentences and that S_1 says *S_2 is not true*, and S_2 says that *S_1 is true*. The L-sentences, S_3 and S_4, are instances of the T-schema.

(S_3) S_2 is true iff S_1 is true

(S_4) S_1 is true iff S_2 is not true.

From (S_3) and (S_4) it follows that S_2 is true iff S_2 is not true, which is a contradiction.

In sum, it is now standard to require that the truth definition for a formal language be given in its metalanguage—an essentially richer language, in order that the truth definition be materially adequate and formally correct. The metalanguage itself may be formalized and studied in a meta-metalanguage. We complete our specification of a full logic by defining model-theoretic and deductive-theoretic consequence relations for M in Chapters 4 and 5, respectively. We shall regard just the sentential connectives, the quantifiers of M, and the identity predicate as logical constants. The language M is a first-order language because its variables range over objects (as opposed to, say, properties). The full logic developed below may be viewed as a version of *classical logic* or a *first-order theory*.

Chapter 4
Model-Theoretic Consequence

In Chapter 2, after highlighting the modal, *a prioricity* and formal features of the common concept of logical consequence, we formulated Tarski's precising definition of logical consequence which attempts to make the common concept of logical consequence more exact.

> A sentence X is a logical consequence of a class K of sentences iff there is no possible interpretation of the non-logical terminology of the language according to which all the K-sentences are true and X is false.

Recall that the Tarskian notion of a possible interpretation is developed in terms of a model: a possible interpretation of the non-logical terminology of the language according to which sentences receive a truth-value is a model of the sentences. This makes possible the move from the precising definition of logical consequence to its characterization in terms of models.

> A sentence X is a logical consequence of a class K of sentences iff all models of K are models of X.

In this chapter, we clarify the notion of a model. We understand models as structures and introduce the notion of *truth in a structure* in terms of which we give a more robust model-theoretic account of logical consequence than the one sketched in Chapter 2. We accomplish this by formalizing the semantics for language M and making the notion of satisfaction precise. The technical machinery to follow is designed to clarify how it is that sentences receive truth-values owing to interpretations of them. We begin by introducing the notion of a structure. Then we revisit the notion of satisfaction, and link structures and satisfaction to model-theoretic consequence. We offer a modernized version of the model-theoretic characterization of logical consequence sketched by Tarski, and so deviate from the details of Tarski's presentation in his (1936) article on logical consequence. Next, we consider the nature of domains and interpretations, the two elements of a structure. This discussion of domains and interpretations will not only help us understand what models depict as representations of logically possible situations, but also set the context for our discussion of logical constants and our evaluation of the model-theoretic consequence characterization of logical consequence.

Truth in a Structure

A structure U for language M is an ordered pair <D, I>. D, a non-empty set of elements, is the domain of discourse, which is the collection of entities that the sentences of M are about. The domain D of a structure for M may be any non-empty set of entities, e.g. the dogs living in Connecticut, the toothbrushes on Earth, the twelve apostles, etc. Interpretation I is a function that assigns to each individual constant of M an element of D, and to each n-place predicate of M a subset of D^n (i.e., a set of n-tuples taken from D). In essence, I interprets the individual constants and predicates of M, linking them to elements and sets of n-tuples of elements from D. In this role, I connects M to the extra-linguistic world. For individual constants *c* and predicates P, the element I_U(c) is the element of D designated by *c* under I_U, and I_U(P) is the set of entities from D assigned by I_U as the extension of P.

By 'structure' we mean an L-structure for some first-order language L. The intended structure for a language L is the course-grained representation of the piece of the world that we intend L to be about. As indicated above, the intended domain D and its subsets that are designated by L's predicates represent the chunk of the world L is being used to talk about and quantify over. The intended interpretation of L's constants and predicates assigns the actual denotations to L's constants and the actual extensions to the predicates. The semantics for the language M of Chapter 3 may be viewed, in part, as an informal portrayal of the intended structure of M, which we refer to as $U^{\mathbf{M}}$. That is, we take M to be a tool for talking about the McKeon family with respect to gender, who is older than whom, who admires whom, etc. To make things formally prim and proper we should represent the interpretation of constants as $I_U{}^{\mathbf{M}}$(matt)=Matt, $I_U{}^{\mathbf{M}}$(beth)=Beth, and so on. And the interpretation of predicates can look like $I_U{}^{\mathbf{M}}$ (Male)= {Matt, Evan}, $I_U{}^{\mathbf{M}}$ (Female)={Beth, Shannon, Kelly, Paige}, and so on. We assume that this has been done.

We are taking a structure U for a language L (i.e., an L-structure) to depict one way that L can be used to talk about a state of affairs. Crudely, the domain D and the subsets recovered from D constitute a rudimentary representation of a state of affairs, and the interpretation of L's predicates and individual constants makes the language L about the relevant state of affairs. Surely, an uninterpreted language can be used to talk about different states of affairs. The class of L-structures represents all the states of affairs that the language L can be used to talk about. For example, consider the following M-structure U'.

D={d∈D| d is a natural number}

$I_{U'}$ (beth)=2	$I_{U'}$ (Male)= {d∈D | d is prime}
$I_{U'}$ (matt)=3	$I_{U'}$ (Female)= {d∈D| d is even}
$I_{U'}$ (shannon)=5	$I_{U'}$ (Parent)=∅
$I_{U'}$ (kelly)=7	$I_{U'}$ (Married)={<d, d'>∈D^2| d+1= d'}
$I_{U'}$ (paige)=11	$I_{U'}$ (Sister)= ∅
$I_{U'}$ (evan)=10	$I_{U'}$ (Brother)={<d, d'>∈D^2| d < d'}
	$I_{U'}$ (OlderThan)={<d, d'>∈D^2| d > d'}
	$I_{U'}$ (Admires)=∅
	$I_{U'}$ (=) = {<d, d'>∈D^2| d=d'}

In specifying the domain D and the values of the interpretation function defined on M's predicates we make use of brace notation, instead of the earlier list notation, to pick out sets. Consider: the sentence

OlderThan(beth, matt)

is true in the intended structure U^{M} for <I_U^{M} (beth), I_U^{M} (matt)> ∈ I_U^{M}(OlderThan). But the sentence is false in U' for <$I_{U'}$(beth), $I_{U'}$(matt)> ∉ $I_{U'}$(OlderThan) (because 2 is not greater than 3). The sentence

(Female(beth) & Male(beth))

is not true in U^{M} but is true in U' for $I_{U'}$ (beth) ∈ $I_{U'}$ (Female) and $I_{U'}$(beth) ∈ $I_{U'}$(Male) (because 2 is an even prime). In order to avoid confusion it is worth highlighting that when we say that the sentence '(Female(beth) & Male(beth))' is true in one structure and false in another we are saying that one and the same wff with no free variables is true in one state of affairs on an interpretation of it and false in another state of affairs on another interpretation of it.

Satisfaction revisited

Note the general strategy of characterizing the truth-conditional property of a sentential connective: the truth of a compound sentence in which it is the dominant connective is determined by the truth of its component well-formed formulas (wffs), which are themselves (simpler) sentences. However, this strategy needs to be altered when it comes to quantificational sentences.

For quantificational sentences are built out of open wffs and, as noted in Chapter 3, these component wffs do not admit of truth and falsity. Therefore, we can't think of the truth of, say,

$$\exists x(\text{Female}(x) \ \& \ \text{OlderThan}(x,\text{paige}))$$

in terms of the truth of 'Female(x) & 'Older than(x,paige)' for some McKeon x. What we need is a truth-relevant property of open formulas that we may appeal to in explaining the truth-value of the compound quantifications formed from them. Tarski is credited with the solution, first hinted at in the following.

> The possibility suggests itself, however, of introducing a more general concept which is applicable to any sentential function [open or closed wff], can be recursively defined, and, when applied to sentences leads us directly to the concept of truth. These requirements are met by the notion of *satisfaction of a given sentential function by given objects*. (1933, p. 189)

The needed property is *satisfaction*. For example, the truth of the above existential quantification will depend on there being an object that satisfies both 'Female(x)' and 'OlderThan(x, paige))'. Earlier we introduced the concept of satisfaction by describing the conditions in which one object satisfies an open formula with one free variable. Now we want to develop a picture of what it means for objects to satisfy a wff with *n* free variables for any $n \geq 0$. We begin by introducing the notion of a *variable assignment*.

A variable assignment is a function g from a set of variables (*g*'s domain) to a set of objects (g's range). With respect to a language L and domain D, a variable assignment assigns elements of D to L's variables. Following Barwise and Etchemendy (2002), we shall say that the variable assignment g is suitable for a wff Ψ of M if every free variable in Ψ is in the domain of g. In order for a variable assignment to satisfy a wff it must be suitable for the formula. For a variable assignment g that is suitable for Ψ, g satisfies Ψ in U iff the object(s) g assigns to the free variable(s) in Ψ satisfy Ψ. Unlike the first-step characterization of satisfaction in Chapter 3, there is no appeal to names for the entities assigned to the variables. This has the advantage of not requiring that new names be added to a language that does not have names for everything in the domain. In specifying a variable assignment g, we write α/v, β/v', χ/v'', … to indicate that $g(v)= \alpha$, $g(v')= \beta$, $g(v'')= \chi$, etc. We understand

$$U \models \Psi[g]$$

to mean that g satisfies Ψ in U.

$$U^M \models \text{OlderThan}(x,y)[\text{Shannon}/x, \text{Paige}/y]$$

This is true: the variable assignment g, identified with [Shannon/x, Paige/y], satisfies 'OlderThan(x,y)' because Shannon is older than Paige.

$$U^M \models \text{Admires}(x,y)[\text{Beth}/x, \text{Matt}/y]$$

This is false for this variable assignment does not satisfy the wff: Beth does not admire Matt.

For any wff Ψ, a suitable variable assignment g and structure U together ensure that the terms in Ψ designate elements in D. The structure U insures that individual constants have referents, and the assignment g insures that any free variables in Ψ get denotations. For any individual constant *c*, c[g] is the element $I_U(c)$. For each variable v, and assignment g whose domain contains v, v[g] is the element g(v). In effect, the variable assignment treats the variable v as a temporary name. Thus, the truth-conditional property of a variable, i.e., what it contributes to the truth conditions of an interpreted sentence in which it occurs, is the element from the domain it designates relative to a variable assignment. For any term t, we define t[g] as 'the element designated by t relative to the assignment g'.

Formalized definition of truth

We now give a definition of truth for the language M *via* the detour through satisfaction. The goal is to define for each formula α of M and each assignment g to the free variables, if any, of α in U what must obtain in order for $U \models \alpha[g]$.

(I) Where R is an *n*-place predicate and $t_1, \ldots, t_n$ are terms, $U \models (t_1, \ldots, t_n)[g]$ iff the *n*-tuple $< t_1[g], \ldots, t_n[g]> \in I_U(R)$

(II) $U \models \sim\alpha[g]$ iff it is not true that $U \models \alpha[g]$

(III) $U \models (\alpha \,\&\, \beta)[g]$ iff $U \models \alpha[g]$ and $U \models \beta[g]$

(IV) $U \models (\alpha \vee \beta)[g]$ iff $U \models \alpha[g]$ or $U \models \beta[g]$

(V) $U \models (\alpha \rightarrow \beta)[g]$ iff either it is not true that $U \models \alpha[g]$ or $U \models \beta[g]$

Before going on to the (VI) and (VII) clauses for quantificational sentences, we introduce the notion of a variable assignment that comes from another variable assignment. Consider the following wff.

$$\exists y(\text{Female}(x) \ \& \ \text{OlderThan}(x,y))$$

We want to say that a variable assignment g satisfies this wff iff there is a variable assignment g′ differing from g at most with regard to the object it assigns to the variable y such that g′ satisfies '(Female(x) & OlderThan(x, y))'. We say that a variable assignment g′ comes from an assignment g when the domain of g′ is that of g and a variable v, and g′ assigns the same values as *g* with the possible exception of the element g′ assigns to v. In general,

$$[g, d/v]$$

refers to an extension g′ of an assignment g by picking out a variable assignment g′ which differs at most from g in that v is in its domain and g'(v)= d, for an element d of the domain D. So, it is true that

$$U^M \models \exists y(\text{Female}(x) \ \& \ \text{OlderThan}(x,y)) \ [\text{Beth}/x]$$

since

$$U^M \models (\text{Female}(x) \ \& \ \text{OlderThan}(x,y)) \ [\text{Beth}/x, \text{Paige}/y].$$

What this says is that the variable assignment Paige/y which comes from Beth/x satisfies '(Female(x) & OlderThan(x,y))' in U^M . This is true, for Beth is a female who is older than Paige.

Now we give the satisfaction clauses for quantificational sentences. Let Ψ be any formula of M in which at most v occurs free in Ψ.

(VI) $U \models \exists v\, \Psi[g]$ iff for at least one element d of D, $U \models \Psi[g, d/v]$
(VII) $U \models \forall v\, \Psi[g]$ iff for all elements d of D, $U \models \Psi[g, d/v]$

Clauses (VI) and (VII) require that each element of the domain is a possible value of a variable, i.e., they require that for any variable v, assignment g, and element d of D, either g(v)=d or *g′*(v)=d for some extension g' of g.

If *α* is a sentence, then it has no free variables and we write $U \models \alpha[g_\varnothing]$, which says that the empty variable assignment satisfies α in U. The empty variable assignment $g_\varnothing$ does not assign objects to any variables. The definition of truth for language M is

A sentence α is true in U iff $U \models \alpha[g_\varnothing]$, i.e. the empty variable assignment satisfies α in **U**.

The truth definition specifies the conditions in which a formula of M is true in a structure by explaining how the semantic properties of any formula of M are determined by its construction from semantically primitive expressions (e.g., predicates, individual constants, and variables) whose semantic properties (i.e., truth-conditional properties) are specified directly. If a set of sentences is true in a structure U we say that U is a model of the set. We now work through some examples. The reader will be aided by referring when needed back to the interpretation of M's predicates given in Chapter 3.

It is true that $U^{\mathbf{M}} \models$ ~Married(kelly, kelly))[$g_\varnothing$], i.e., by (II), it is not true that $U^{\mathbf{M}} \models$ Married(kelly,kelly))[$g_\varnothing$], because < kelly[$g_\varnothing$], kelly[$g_\varnothing$]> $\notin$ $I_U^{\mathbf{M}}$(Married). Hence, by (IV)

$U^{\mathbf{M}} \models$ (Married(shannon, kelly) v ~Married(kelly, kelly))[$g_\varnothing$] by (IV).

Our truth definition should confirm that

$$\exists x \exists y \text{Admires}(x,y)$$

is true in U^M. Note that by (VI) $U^M \models$ ∃yAdmires(x,y)[$g_\varnothing$, Paige/x)] since $U^M \models$ Admires(x,y)[[$g_\varnothing$, Paige/x)], kelly/y]. Hence, by (VI)

$$U^M \models \exists x \exists y \text{Admires}(x,y)[g_\varnothing].$$

The sentence

$$\forall x \exists y(\text{Older}(y,x) \rightarrow \text{Admires}(x,y))$$

is true in U^M. By (VII), $U^M \models$ ∀x∃y(Older(y,x)→Admires(x,y))[$g_\varnothing$] iff for all elements *d* of D, $U^M \models$ ∃y(Older(y,x)→Admires(x,y)) [$g_\varnothing$, d/x]. This is true. For each element *d* and assignment [$g_\varnothing$, d/x], $\mathbf{U^M} \models$ (Older(y,x)→Admires(x,y)) [[$g_\varnothing$, d/x], d'/y] i.e., there is some element d' and variable assignment g differing from [$g_\varnothing$, d/x] only in assigning d' to *y*, such that g satisfies '(Older(y,x)→Admires(x,y))' in U^M.

Model-Theoretic Consequence Defined

For any set K of M-sentences and M-sentence X, we write

$$K \models X$$

to mean that every M-structure that is a model of K is also a model of X, i.e., X is a model-theoretic consequence of K.

(1) OlderThan(paige, matt)

(2) ∀x(Male(x)→OlderThan(paige,x))

Note that both (1) and (2) are false in the intended structure U^M. We show that (2) is not a model-theoretic consequence of (1) by describing a structure which is a model of (1) but not (2). The above structure U' will do the trick. By (I) It is true that U' ⊨ OlderThan(paige, matt) [$g_\varnothing$] because <(paige) $g_\varnothing$, (matt) $g_\varnothing$> ∈ $I_{U'}$ (OlderThan) (because 11 is greater than 3). But, by (VII), it is not the case that

U' ⊨∀x(Male(x)→OlderThan(paige,x)) [$g_\varnothing$]

since [$g_\varnothing$, 13/x] does not satisfy '(Male(x)→OlderThan(paige,x))' in U' according to (V) for U' ⊨ Male(x) [$g_\varnothing$, 13/x] but not U' ⊨ OlderThan(paige,x)) [$g_\varnothing$, 13/x] . So, (2) is not a model-theoretic consequence of (1). Consider the following sentences.

(1) (Admires(evan, paige)→Admires(paige,kelly))
(2) (Admires(paige, kelly)→Admires(kelly, beth))
(3) (Admires(evan, paige)→ Admires(kelly, beth))

(3) is a model-theoretic consequence of (1) and (2). For assume otherwise. That is, assume that there is a structure U" such that

(i) U" ⊨ (Admires(evan, paige)→Admires(paige,kelly)) [$g_\varnothing$]

and

(ii) U" ⊨ (Admires(paige, kelly)→Admires(kelly, beth)) [$g_\varnothing$]

but not

(iii) U" ⊨ (Admires(evan, paige)→ Admires(kelly, beth)) [$g_\varnothing$].

By (V), from the assumption that U" does not satisfy (iii) it follows that U" ⊨ Admires(evan, paige) [$g_\varnothing$] and not U" ⊨ Admires(kelly, beth) [$g_\varnothing$]. Given the former, in order for (i) to hold according to (V) it must be the case that U" ⊨Admires(paige,kelly))[$g_\varnothing$]. But then U" ⊨Admires(paige,kelly)) [$g_\varnothing$] and not U" ⊨ Admires(kelly, beth) [$g_\varnothing$], which, again appealing to (V), contradicts our assumption (ii). Hence, there is no such U", and so (3) is a model-theoretic consequence of (1) and (2).

Here are some more examples of the model-theoretic consequence relation in action.

(1) ∃xMale(x)

(2) ∃xBrother(x,shannon)
(3) ∃x(Male(x) & Brother(x,shannon))

(3) is not a model-theoretic consequence of (1) and (2). Consider the following structure U.

D={1, 2, 3}
For all M-individual constants c, I_U(c)=1.
I_U (Male)={2}, I_U (Brother)={<3,1>} For all other M-predicates P, I_U(P)=∅.

Appealing to the satisfaction clauses (I), (III), and (VI), it is fairly straightforward to see that the structure U is a model of (1) and (2) but not of (3). For example, U is not a model of (3) for there is no element d of D and assignment [d/x] such that U ⊨ (Male(x) & Brother(x, shannon))[$g_\varnothing$, d/x].

Consider the following two sentences.

(1) Female(shannon)
(2) ∃xFemale(x)

(2) is a model-theoretic consequence of (1). For an arbitrary M-structure U, if U ⊨ Female(shannon) $g_\varnothing$, then by satisfaction clause (I), shannon[$g_\varnothing$] ∈ I_U (Female), and so there is at least one element of D, shannon[$g_\varnothing$], in I_U (Female). Consequently, by (VI), U ⊨ ∃xFemale(x) [$g_\varnothing$].

For a sentence α of M, we write

$$\models \alpha$$

to mean that α is a model-theoretic consequence of the empty set of sentences. This means that every M-structure is a model of α. Such sentences represent logical truths; it is not logically possible for them to be false. For example,

$$\models (\forall x\text{Male}(x) \rightarrow \exists x\text{Male}(x))$$

is true. Here's one explanation why. Let U be an arbitrary M-structure. We now show that

$$U \models (\forall x\text{Male}(x) \rightarrow \exists x\text{Male}(x))\ [g_\varnothing].$$

If U ⊨(∀xMale(x) [$g_\varnothing$] holds, then by (VII) for every element d of the domain D, U ⊨Male(x) [$g_\varnothing$, d/x]. But we know that D is non-empty, by the requirements on structures, and so D has at least one element d. Hence for at least one element d of D, U ⊨Male(x) [$g_\varnothing$, d/x], i.e., by (VI), U ⊨ ∃xMale(x))

[$g_{\varnothing}$]. So, if U $\models$(∀xMale(x) [$g_{\varnothing}$] then U $\models$ ∃xMale(x)) [$g_{\varnothing}$]. Therefore, according to (V),

$$U \models (\forall x Male(x) \rightarrow \exists x Male(x)) [g_{\varnothing}].$$

Since U is arbitrary, this establishes

$$\models (\forall x Male(x) \rightarrow \exists x Male(x)).$$

If we treat '=' as a logical constant and require that for all M-structures U, I_U (=) = {<d, d'>∈ D^2| d=d'}, then M-sentences asserting that identity is reflexive, symmetrical, and transitive are true in every M-structure, i.e. the following hold.

$\models\forall x(x=x)$
$\models\forall x\forall y((x=y)\rightarrow(y=x))$
$\models\forall x\forall y\forall z((x=y)\&(y=z))\rightarrow(x=z))$

Structures which assign {<d, d'>∈ D^2| d=d'} to the identity symbol are sometimes called normal models. Letting 'Ψ(v)' be any wff in which just variable v occurs free,

$$\forall x\forall y((x=y)\rightarrow(\Psi(x)\rightarrow \Psi(y))$$

represents the claim that identicals are indiscernible—if x=y then whatever holds of x holds of y—and it is true in every M-structure U that is a normal model. Treating '=' as a logical constant (which is standard) requires that we restrict the class of M-structures appealed to in the above model-theoretic definition of logical consequence to those which are normal models.

We now sharpen our understanding of a structure by elaborating on the nature of its two elements, domains and interpretations. As illustrated above, a structure interprets a language L by assigning a truth-value to each of L's sentences, thereby representing a possible use for L, i.e., representing one way that L can be used to talk about a state of affairs. Recall that a domain D and its subsets constitute a rudimentary representation of a state of affairs. An assignment to L's variables, and an interpretation of L's predicates and individual constants makes the relevant state of affairs what L is about. The fact that a set K of sentences are true and sentence X is false relative to a structure is evidence that X is not a logical consequence of the K-sentences, because we regard *the possible use for L* depicted by the structure as demonstrating the logical possibility of the K-sentences being true and X being false.

We understand logical possibility in terms of a possible use for some language. For example, we understand ‘it is logically possible that a set K of L-sentences be true’ as there is a possible use for L according to which they are all true. By virtue of depicting possible uses for language, structures are representations of logical possibilities. The class of structures for a language L represents the class of logically possible situations we appeal to in determining logical truth and consequence in L. One question that emerges is: How should we conceive of domains and interpretations in order for structures to depict possible uses of language and thereby represent logical possibilities?

Our response begins by clarifying what qualifies as a domain. Getting clear on which collections may serve as domains will enable us to clearly see what work a domain does in the use of a structure to represent logical possibilities for sentences from a first-order language. Then we clarify constraints on admissible interpretations, which shall shed light on, among other things, how the notion of semantic modality (i.e., the possibility of a meaningful expression having an alternative meaning) figures in the model-theoretic conception of logical consequence. Our discussion of domains and interpretations will highlight the significance of structures in the model-theoretic characterization of logical consequence and set the context for our evaluation of the later.

The Metaphysics of Domains

> Logicians usually merely say that, in giving an interpretation, we need to specify a non-empty domain, without bothering themselves with the means by which this is to be done. (Sometimes they add that the domain must be a set, and, in this, a severe restriction may be implicit: but to make it explicit we should have to go very deeply into the notion of a set. (Dummett (1991a), p. 312)

Instead of clarifying domains by doing the metaphysics of sets, we pursue a different strategy (motivated by Dummett). The elements of a domain used to interpret a language L are distinguished (either implicitly or explicitly) in L’s metalanguage *via* what we may call (following Forbes ((1994), p.151) a *domain indicator*. The domain indicator is a metalinguistic predicate that designates a property which determines a domain over which L’s variables range.

We explicitly use a domain indicator to specify a domain when we produce a token of the predicate that designates a property whose extension is the domain. An implicit use of a domain indicator is its specification by a list of the names of its elements. That a name successfully refers to an element of the domain requires that the element possess a unique characterizing feature. On the Fregean way of spelling this out, each of the names on the list has a sense which includes a specification of the kind of object named. Thus a domain given in terms of a list of names will yield a disjunctive property which determines the membership condition of the domain, and, therefore, the corresponding disjunctive predicate serves as a domain indicator. It is not assumed that there be a non-disjunctive property true of every element of a domain. A list of names—a, b, c, d—may determine a domain even though the names refer respectively to an ant, Michigan, Barack Obama, and π, which are of radically different kinds. We leave open the question for now whether it is possible to characterize a domain with an infinite list of names that corresponds to an infinite disjunctive property. We pick up this issue later when we consider the possibility of there being a domain of everything. In explaining domain indicators, we'll focus on domain indicators that designate non-disjunctive properties.

Our strategy is to clarify the nature of domains by identifying the types of predicates that may serve as domain indicators. Succinctly, a domain indicator must designate a sortal property that is collectivizing. A sortal property is a property that supplies criteria of application, identity, and individuation. A collectivizing property is a property for which there is a collection whose members are just the objects that have it (Potter (2004), p.25). The idea here is that it is only collectivizing, sortal properties that determine domains. The requirement that domain indicators designate collectivizing properties is robust. As we shall highlight below, there are sortals that are non-collectivizing. Furthermore, a property that is either indefinitely extensible or true of every object that possesses an indefinitely extensible property is non-collectivizing, and so does not determine a domain. We now quickly sketch the rationale for requiring that a domain indicator designate a sortal property that is collectivizing. We'll elaborate later as we clarify what sortals and indefinitely extensible properties are. For ease of exposition we call a property hazy if it is either a non-sortal or a non-collectivizing property.

Intuitively, the concept of logical possibility applies to truth-bearers' possession of particular truth-values, whether these be *true*, *false*, or whatever. For us, truth-bearers are sentences, and so what is or isn't logically possible are sentences possessing particular truth-values. Crudely, a possible use of a language is one way of assigning truth-values to all of its sentences. We regard such a use as a representation of a logical possibility according to which the sentences possess the truth-values thus assigned. Every possible use of a language represents a logical possibility. Different understandings of *possible use* motivate different logics. For example, inspired by Tarski, we take a possible use of a language as one way of assigning the values *true* or-exclusive *false* to each sentence. This yields a two-valued logic, according to which *false* means *not true*. Alternatively, we could take a possible use as one way of assigning, say, the values, *true*, *false*, or *neither true nor false* to each sentence. Accounting for such uses of language yields a three-valued logic, according to which a sentence can fail to be true in one of two ways.

In order for a structure to depict a possible use for a language and thereby represent a logically possible situation, it must interpret the language by assigning a value from a set of designated truth-values to each sentence. If what a hazy property is true of were to serve as a domain for a structure, then that structure would not assign a determinate truth-value to every sentence of the relevant language. It would not depict a possible use for the language, and would not, therefore, represent a logically possible situation. But then such a structure would not be relevant to determining logical truth and consequence. In order for a structure to represent a logically possible situation, its domain must be the extension of a sortal property that is collectivizing.

In sum, the function of a domain indicator is to determine what counts as an individual (i.e., what counts as an object) for purposes of assigning truth-values to *all* sentences from the language we aim to interpret. No hazy property determines what counts as an individual for purposes of assigning truth-values to *all* sentences from the language. Hence, a domain indicator must designate a collectivizing, sortal property. We now further explain our characterization of domain indicators and elaborate on its motivation as sketched in the previous paragraph.

A domain indicator must designate a property that supplies a definite criterion of application

A property supplies a definite criterion of application if it is determinate what is to hold good of an object for it to fall under it (Dummett (1981) pp.

74-76, 233-234, 546-548, *et al.*). We say that a property is vague if it lacks a definite criterion of application, i.e., there are things, borderline cases, for which it is indeterminate whether the property holds for them or not. There are two approaches to accounting for such indeterminacy (we borrow from Sainsbury (1995), Ch. 2).

The metaphysical (semantic) approach to vagueness: vagueness is an objective feature of reality. In particular, there are properties that are intrinsically vague. For such a property there are objects for which there is no definite fact of the matter whether the property holds for them.

The epistemological approach to vagueness: vagueness is nothing but ignorance. It is not an objective feature of reality. For every property and every object there is a fact of the matter whether the property is true of the object, though this may be a fact that we are necessarily ignorant of.

Each approach offers a unique explanation why vague properties should be ruled out as domain indicators. We now elaborate.

Recall that the function of a domain indicator is to determine what counts as an individual for purposes of assigning truth-values to *all* sentences from the relevant language. If the extension of a vague property were to serve as a domain of a structure U, then there will be sentences that U fails to assign a truth-value. Hence, a vague property cannot fulfill the function of a domain indicator. To see this, suppose we have a property P such that it is true of just α, but it is indeterminate whether it is true of β. If we were to treat P's hazy extension as a domain for a structure **U**, then it would be indeterminate whether or not

$$\mathbf{U} \models \exists x \exists y\, (x \neq y)\ [g_{\varnothing}],$$

and indeterminate whether or not there is an interpretation I_U such that

$$\mathbf{U} \models \text{Male(evan)\&\textasciitilde Male(paige)}\ [g_{\varnothing}].$$

If we adopt the metaphysical approach and assume that P's indeterminacy is metaphysical, then there is no definite fact of the matter about whether the two sentences are true relative to U. This impacts the model-theoretic determination of logical truth and consequence. For example, letting α be any M-sentence, the following hold.

$\models (\alpha \vee \sim \alpha)[g_{\varnothing}]$
$\models \sim (\alpha \;\&\!\sim \alpha)[g_{\varnothing}]$

The first reflects what is sometimes called the law of excluded middle, and the second the law of non-contradiction. Obviously, both presuppose that structures assign a truth-value to every sentence. If we let the indeterminate extension of a vague property serve as a domain for a structure U that interprets language M, then both laws fail, because there will be M-sentences that are instances of the laws but are not evaluable in U, and, therefore, not true in U.

On the epistemological approach to vagueness, there is a definite fact of the matter regarding whether β is a P. If P's indeterminacy is epistemological, then a structure U whose domain is P's extension assigns $\exists x\exists y\ (x\neq y)$ a truth-value, even if we may be unable to determine it. Furthermore, there is a fact of the matter whether there is an interpretation that makes Male(evan) &~Male(paige) true relative to such a U. On the epistemological approach, admitting vague properties as domain indicators does not alter the model-theoretic determination of logical truth and consequence (e.g., the laws of excluded middle and non-contradiction are preserved). Here the motivation for restricting domain indicators to those properties that provide a definite criterion of application is the desire for a structure to be useful in determining logical truth and consequence. Obviously, structures with epistemically indeterminate domains are not useful in determining the model-theoretic consequence relation. Hence, in order to insure the epistemic accessibility of *truth in a structure* vague properties should be prohibited from being domain indicators.

In sum, we require that a domain indicator supply a definite criterion of application to so that structures assign a determinable truth-value to each M sentence and represent epistemically accessible logical possibilities. Note well: the requirement that a structure assign a determinable truth-value to each sentence does not determine which truth-values it assigns or how many there are. For example, we may define a vague monadic-property P on a domain D as follows. One subset of D serves as the property's extension, i.e., the set E_P of objects the property is true of. Another subset of D serves as the property's anti-extension, i.e., the set $\bar{E}_P$ of objects that P is false of. A third subset of D serves as P's penumbra, the set $\hat{E}_P$ of objects for which it is indeterminate whether P is true of them. Let R be a 1-place predicate, and *t* a term. Consider a structure U according to which $I_U(R) = E_P$. A variable assignment g satisfies R(t) relative to U iff $t[g] \in E_P$; g does not satisfy R(t) relative to U iff $t[g] \in \bar{E}_P$. Assignment g neither satisfies nor does not satisfy R(t) relative to U iff $t[g] \in \hat{E}_P$. If instances of the law of excluded middle or non-contradiction are not true in such a structure U, it is because they have the truth-value of *neither true nor false* and not because they are not evaluable in U. Treating *neither true nor false* as a truth-value rather than as a truth-value gap is not idiosyncratic (e.g., see Priest (2001) p. 119, and for brief discussion see Haack (1996) pp. 55-580). Certainly, the requirement that a domain indicator supply a definite criterion of application is not grounds for restricting the states of affairs that a language may be about to ones that are bivalent.

A domain indicator must designate a property that supplies a definite criterion of identity

A property P supplies a definite criterion of identity when it is determinate what is to count as one and the same P. If asked to distinguish the animals in Potter Park Zoo from the benches, one can do so by employing the criterion of application for *animals*, which determines when it is correct to say 'That is an animal'. However, if asked to count how many animals there are at the zoo, one cannot comply unless 'animals' is first disambiguated. Is one to count species or individual beasts? Counting involves identity, because in order to count correctly the same object must not be counted twice. Different criteria of identity are associated with animal species and beasts. More generally, we use different criteria of identity for objects of different kinds. The above examples show that a criterion of identity for P is not derivable from a criterion of application for P (Dummett (1981), pp. 74-75, 233, 573). The second request is ambiguous in a way that the first is not. A criterion for determining when statements of the form 'That is the same animal as the one which …' are true is not derivable from a criterion for determining the correctness of saying, 'That is an animal'.

Many common nouns, typically called *count nouns*, designate properties that yield a principle of identity (e.g., 'animal', 'person', 'book', and' desk'). Almost all adjectives (e.g., blue, smooth, slimy, heavy, etc.) designate properties associated with a criterion of application but not a criterion of identity. Hence, there is no definite range of objects to which these properties belong. For example, the question, 'What is *this* smoothness?", lacks sense because there is no criterion of identity associated with a noun such as 'smoothness' (Dummett (1981), p.77). As Wright ((1983), pp.2-3) points out, a full understanding of what it is for something to be a person requires not only an understanding of its criterion of application, but also an understanding of the states of affairs that make sentences of the form *x is the same person as y* true. This requires knowing both what it is to encounter the same person again and what constitutes personal survival. However, in order to know what makes sentences of the form *x is the same smooth thing as y* true, one needs first to know what kind of things x and y are, and understanding 'smooth' does not require that knowledge. Understanding what smoothness is requires an ability to properly employ its criterion of application, but it does not require an ability to re-identify smooth things.

Typically, a criterion of identity for property P is an informative and non-circular principle of the form: For all Ps x and y, x is identical with y iff x and y stand in a relation R to one another (Lowe (1997), p.620). Relation R may be specified with reference to the identity of things of a kind other than P, but in order to avoid circularity R must specifiable without reference to the identity of the Ps. For example, sets x and y are identical iff x and y have the same members. Real numbers x and y are identical iff x and y are greater than or less than the same rational numbers. According to a Quinean understanding of events, events x and y are identical iff x and y are the total contents of the same spatio-temporal region r. Note that we appeal to the identity criteria of set-elements, rationals, and spatio-temporal regions, respectively. It is debatable whether an adequate criterion of identity can be given that does not appeal to the identity of things of any kind. Lowe argues that this is unlikely and presses the implication that, "in any system of ontology objects of a certain kind must be included for which no informative and non-circular identity criterion can be supplied" (Lowe (2003), p. 91). This raises the issue of how we are to understand quantification over such objects. Unfortunately, we cannot pursue this here. Why require that a domain indicator supply a criterion of identity? We split a rationale into four reasons.

(i) Recall that the job of a domain indicator in interpreting a language is to determine what counts as an object for purposes of assigning truth-values to the language's sentences. Following Frege, in order to have a conception of any particular (self-subsistent) *object*, we have to know what constitutes the correct recognition of that object as the same again (Frege (1950) sections 56 and 62, Dummett (1981) p. 578). So, if a predicate (or a range of them) is introduced to fix a domain, a grasp of the appropriate criterion of identity or the range of different ones in use constitutes the basis of an apprehension of any element of the domain. Such apprehension is needed in order to understand reference for a term. For example, variable assignments presuppose the means for fixing reference for variables, and the latter requires that we can distinguish one referent from another (for discussion, see Epstein (2001), pp.81-84, Dummett (1991b), p. 326).

(ii) Informally, assertions about identity employ implicit criteria regarding what properties count and what properties are irrelevant to the truth of the assertion. For example, implicit in the assertion that 'A' is the same letter as 'A' is that it's letter-type that counts and we should ignore that the letters are different tokens. Hence, the truth of the assertion turns on the criterion of identity associated with letter-type. For any structure U, U $\models$ (x=y)[g] iff variable assignment g assigns the same object of the domain to both variables. Essentially 'x=y' is interpreted as x[g] and y[g] are the same P, where P is a domain indicator. But to distinguish objects in terms of P makes sense only if P supplies a criterion of identity. The very intelligibility of the interpretation of '=' in a domain requires that the domain be the extension of a property that supplies a criterion of identity.

(iii) We cannot assign a truth-value to a sentence such as '∃x(Male(x) & Brother(x, shannon))' unless we understand what it means to assign the same P to the occurrences of the variable in '(Male(x) & Brother(x, shannon))', where P is the domain indicator.

(iv) Interpreted predicates designate their extensions, which are subsets of some previously determined domain of objects. So, a determinate interpretation of a predicate presupposes a prior specification of a domain. Any means of specifying a domain for a structure U which did not supply a criterion of identity, for instance by saying that it was to consist of green and white things in East Lansing, MI on November 13th, 2009 would not tell us what would count as a determinate interpretation of any one-place predicate over this domain. For example, the admissibility of interpretation according to which 'Male(matt)' and '~Male(beth)' are true turns on I_U(matt) not being the same green and white thing in East Lansing at the above time as I_U (beth). But this is indeterminate, because the property lacks criteria of identity.

With respect to our language M, a domain for an M-structure is specified prior to the interpretation of M's predicates. Hence, the identity relation over a domain used to interpret M is determined by the ontology of M's metalanguage and not in terms of the interpreted vocabulary of M. This determination is "the source of the finest possible identity relation" (Dummett (1991b), p.315) over the domain. To illustrate, consider a language M′, which is just like M except that it lacks '='. That $\alpha \in$ D and $\beta \in$ D are indistinguishable relative to the interpreted predicates of M′ (i.e., for any atomic wff Φ, α satisfies Φ iff β satisfies Φ) is not sufficient for α and β being one and the same object for they may be distinguishable in the metalanguage.

A domain indicator must designate a property that supplies a criterion of individuation

A principle of individuation tells us what is to count as one instance of a given property. It is a general principle for marking off individual instances of a property. More specifically, a property supplies a principle of individuation iff both (i) and (ii) obtain (we follow Koslicki (1997)).

(i) The property isolates what falls under it in a definite manner; it divides up what it is true of into discrete units.

(ii) The property does not permit any arbitrary division of what it is true of into parts.

(i) insures that the property delineates its extension into discrete units. What the property is true of is not undifferentiated stuff; it draws precise boundaries around each thing in its extension. (ii) concerns the internal structure of these units. Once we are down to the level of discrete units, (ii) tells us that we cannot go on dividing the original units and expect to find more units of the same kind. Instances of a property P that lack intrinsic unity can only be singled out in *ad hoc* ways, of which there are indefinitely many. We may construe "*ad hoc* ways of singling out instances" as "a myriad of unprincipled divisions" (Koslicki (1997), p. 424). Randomly picked proper parts of stuff S are still S. Quantities of stuff can be divided and combined in *ad hoc* ways, while remaining quantities of the same stuff; objects cannot be.

If a property supplies a principle of individuation, then it divides what it is true of in such a way that its instances can be referred to. Borrowing from Quine, common nouns used to designate such properties (e.g., 'apple', 'tree', 'planet') have built in modes of dividing their reference (Quine (1960), p.91ff). Mass nouns such as 'snow', 'water', 'gold', 'mud', and 'air', which stand for undifferentiated stuff (or substance), are not associated with a principle of individuation and consequently have the semantic property of referring cumulatively (e.g., any sum of parts which are water is water) and do not divide their reference (Quine (1960), p.91). Consider an example from Quine: 'shoe', 'pair of shoes', and 'footwear' all range over the same stuff, but the first two divide their reference differently and the third not at all. Of course, stuff can be "individualized" (from Epstein (2001, p.156) by various means, e.g., a ball of snow, a cup of water, an ounce of gold. The Quinean point here is that the logical force of individuation is reference dividing. Therefore, the two truths

Snow is white
Water is a fluid

cannot be represented respectively as,

$\forall$x (Snow(x) $\rightarrow$ White (x))
$\forall$x(Water(x) $\rightarrow$Fluid(x))

because 'Snow' and 'Water' do not divide up what they are true of into discrete units that variables can refer to under an assignment. This illustrates why mass nouns cannot be employed as domain indicators without reconfiguring them so that the resulting predicate provides a principle of individuation resulting in the individualization of what they are true of.

A principle of individuation is not derivable from a principle of identity (Lowe (1997), p. 616). Mass terms such has 'gold' and 'water' convey identity criteria (e.g., it is sensible to ask if the gold on the table (scattered about in the form of dust) is the same that formerly composed a certain wedding ring), even though it makes no sense to ask how many portions of gold are on the table. Of course, it does make sense to ask how much gold is on the table. Thus, sentences, '$\exists$x$\exists$y (x$\neq$y)' and '$\exists$x$\forall$y(x=)' are evaluable relative to structure U only if U's domain is countable and this requires that it be the extension of a property with a principle of individuation, as well as a principle of identity. Frege notes in effect ((1950), p. 66) that "being a red thing" does not determine a finite number because it fails to supply a criterion of individuation. The trouble about counting the red things in the room is not that you cannot make an end to counting them but that you cannot make a beginning (Geach (1962) pp.38-39). The question, how many angels can dance on a head of a pin?, is senseless unless there is a principle of individuation for angels.

A property P that supplies definite criteria of application, identity, and individuation is called a sortal property because it sorts what it is true of into a definite kind. The requirement that a domain indicator designate a sortal property restricts domains to those collections that are extensions of sortal properties. Since a domain indicator fixes what counts as an object on a given interpretation of M, we may derive a partial model-theoretic conception of what the term 'object' means in the most general sense of that term. Succinctly, from the perspective of first-order model-theoretic semantics, 'object' properly applies to any item that can be the referent of an individual constant or, equivalently, the value of a variable under an assignment (an object can be made the referent of an individual constant iff it can be made the value of a variable). Consequently, if the term 'object' refers to a thing, then it is an instance of a sortal property, since no thing can be made the value of a variable or the referent of an individual constant unless it is an instance of a sortal property. Whether we can quantify over, say, clouds, waves, or dreams depend, in part, on whether these things qualify as "objects" in the above sense. So, our above use of a structure with a domain of natural numbers commits us to the thesis that natural numbers are objects.

We have been highlighting the general concept of an object required by model-theoretic semantics. However, the particular kind of objects that make up a domain is irrelevant to fixing the model-theoretic consequence relation. Relative to M, two structures U and U', are *isomorphic* iff there is a one-to-one correspondence between D and D' where each element d of D is paired off with a unique element d' of D', i.e., $f(d_1) = d'_1, \ldots, f(d_n) = d'_n$ (thus D and D' have the same cardinality) such that the following holds.

For each individual constant c of M, $I_U(c) = d$ iff $I_U'(c) = d'$.
$<t_1[g], \ldots, t_n[g]> \in I_U(R)$ iff $<t_1[g'], \ldots, t_n[g']> \in I_U'(R)$.

If U and U' are isomorphic models of a first-order language L, then for any L-sentence α, $\mathbf{U} \models \alpha[g_\varnothing]$ iff $\mathbf{U}' \models \alpha[g_\varnothing]$. Let's call this (following Shapiro) the isomorphism property of structures. Isomorphic models U and U' are indistinguishable relative to the L-sentences they make true even if the objects in D and D' are of different kinds. If $\mathbf{U} \models \alpha[g_\varnothing]$ and D consists of three frogs, then we can construct an interpretation I′ such that $\mathbf{U}' \models \alpha[g_\varnothing]$ where D′ consists of, say, three paper cups. From the isomorphism property it follows that the specific kind of the objects in a domain doesn't matter for model-theoretic consequence. The size of a domain (i.e., its cardinality) does matter, e.g., for a structure U, $\mathbf{U} \models$ (Female(shannon) & ~Female(matt))$[g_\varnothing]$ only if D has at least two elements. So, for any set D of objects and any L-sentence α, whether or not there exists an interpretation I of L such that $\mathbf{U} \models \alpha[g_\varnothing]$ turns on how many objects are in D and not at all on the kind of objects. The significance of the specification of a domain as, say, the McKeons or certain natural numbers in characterizing a structure for a first-order language such as M is clarified by the rationale for requiring that a domain indicator designate a sortal property.

A domain indicator must designate a property that is collectivizing

According to Bertrand Russell, a property is self-reproductive (equivalently, indefinitely extensible) if

> …given any class of terms all having [the] property, we can always find a new term also having the property in question. Hence we can never collect *all* of the terms having the said property into a whole; because whenever we hope to have them all, the collection which we have immediately proceeds to generate a new term also having the said property. ((1906), p. 144)

Dummett writes that,

> A concept is indefinitely extensible if, for any definite characterization of it, there is a natural extension of this characterization, which yields a more inclusive concept; this extension will be made according to some general principle for generating such extensions, and, typically the extended characterization will be formulated by reference to the previous, unextended, characterization ((1963), p.195-196).

Working off Potter ((2004), p. 29), a property P is indefinitely extensible if there is a *principle of extendibility* associated with P, which when applied to some Ps generates an object that is not among them but is nevertheless a P. A property is *collectivizing* only if there is a collection whose members are just the objects which have it (Potter (2004), p. 25). A domain is a collection. So, if a property is non-collectivizing, then what it is true of cannot be a domain. Since indefinitely extensible properties are non-collectivizing, no domain indicator designates an indefinitely extensible property. We now illustrate the notion of indefinite extensibility with two examples (borrowing from Boolos (1993) and Clark (1993)).

The elements of the following progression are *ordinal numbers*.

$0, 1, 2, 3, \ldots\omega, \omega+1, \omega+2, \omega+3, \ldots 2\omega, 2\omega+1, 2\omega+2, 2\omega+3, \ldots 3\omega, 3\omega+1, 3\omega+2, 3\omega+3, \ldots 4\omega\ldots$

Let's call this *the ordinal progression*. Principles (1) and (2) characterize the ordinal progression.

(1) For all ordinals α, there exists an ordinal $\alpha+1$ which is the immediate successor of α.

(2) For any definite unbounded succession of ordinals, there is a least ordinal greater than all those numbers, i.e., there exists an ordinal that is the limit of those numbers. In other words: for any definite succession of defined ordinals for which there is no largest, there exists a new ordinal which is the limit of those numbers, i.e., it is the next number larger than all of them.

Let each ordinal be identified with the collection of preceding ordinals. E.g., $3 = \{0,1,2\}$, $\omega = \{1, 2, 3, \ldots\}$. Principle (2) characterizes the limit ordinals, which have no immediate predecessors i.e., for a limit ordinal α, there is no ordinal number β such that $\beta + 1 = \alpha$. The first limit ordinal is ω, and then 2ω, $3\omega \ldots \omega^2$, and so on.

Is the ordinal progression the collection of all ordinal numbers? Answer: No. Suppose that it is. Let's refer to it as E_0. Then, by (2), there exists its limit: ON. It is the least ordinal greater than every ordinal in E_0. Hence, there is another collection of ordinals, E_1, that consists of E_0 plus ON, which in turn can be extended, by (1), to include ON +1, ON +2, and so on. Applied to any E_n, the principle of extendibility for *ordinal number*, construed as the operation $E_n | \rightarrow E_n \cup \{E_n\}$ (Boolos (1993), p.215), generates another collection of ordinals that includes one new to E_n, i.e., it includes $\{E_n\}$, the limit of E_n. The property of *being an ordinal number* is indefinitely extensible and is, therefore, non-collectivizing.

Consider the sortal property of *being a domain*. Is there a collection C that contains every domain? If so, then given that *being a domain* is a sortal property, C is a domain containing all domains. We now demonstrate that there is no such domain C, and, therefore, no collection of every domain. The argument shows why being a domain is indefinitely extensible.

We appeal to the subdomain principle: if the extension E of a property P is a subset of a domain, then E itself is a domain. In other words: for any domain D, if $E \supseteq D$, then E is a domain (and a predicate that designates E is a domain indicator). Suppose that there is a domain C consisting of every domain. Consider the following property P with extension E that is defined on C as follows: (1) for all domains $D \in C$, $D \in E$ iff $D \notin D$. From our hypothesis that C is a domain, it follows that E is a domain by the subdomain principle. Hence, (2) $E \in C$. From (1) and (2), it follows that (3) $E \in E$ iff $E \notin E$, which is a contradiction. Therefore our initial assumption is false: C is not a domain of all domains for it doesn't contain domain E. This illustrates that the property of *being a domain* is an indefinitely extensible property. It is a non-collectivizing property. The principle of extendibility associated with *being a domain* may be formulated as the operation: $E_n \mapsto E_n \cup \{\{y \in E_n: y \notin y\}\}$ (borrowing from Boolos (1993), p. 215).

We've said that from the perspective of first-order model-theoretic semantics, a necessary condition for the proper application of the term 'object' to an item is that it is an instance of a sortal property. This is not sufficient. There are sortal properties such as *being a domain* and *being an ordinal number* that cannot be used to determine what counts as an object on any interpretation of a first-order language, because they are indefinitely extensible properties, and, therefore, they are non-collectivizing. Hence, by the lights of first-order model-theoretic semantics, things *qua* instances of *being a domain* or *qua* instances of *being an ordinal number* do not qualify as objects. To avoid misunderstanding, it is worth making some general remarks regarding sortals that are indefinitely extensible, and, therefore, non-collectivizing.

Let P be an indefinitely extensible sortal property. Then there is a series of extensions

$$E_0, E_1, E_2, E_3, E_4, \ldots$$

and principle of extendibility, associated with P, that takes us from one of these extensions to the next in the sequence. The above sequence corresponds to the sequence of properties

$$P_0, P_1, P_2, P_3, P_4, \ldots$$

For each *n*, E_n is the extension of sortal property P_n. The P_i are determinate specifications of P, and the corresponding E_i are domains. Model-theoretically, 'object' may apply to an instance of P_n for any *n.* What is the status of P? Following Dummett ((1963), p. 195-196, (1981) p. 476, pp.532-533, (1991b), p. 316), it is the *intuitive property* of being a P from which we recover the associated principle of extendibility. Property P may be determinate insofar as it is a sortal, but it is not collectivizing, and so it does not determine a domain.

Following Lowe ((1997), p. 616-618), there are two possible answers to the question, what does 'object' mean in the most general sense of that term?

The linguistic answer: the term 'object' properly applies to anything that can be quantified over. Quine's dictum, "to be is to be the value of a variable."

The metaphysical answer: 'object' properly applies to anything that is an instance of a sortal property.

Since, there are sortals that are indefinitely extensible and no indefinitely extensible property determines a domain, these two answers really *are* different. The model-theoretic semantic conception of what the term 'object' (in its most general sense) applies to is in accordance with the linguistic answer.

A domain of a structure is a non-empty collection of things that the structure interprets some first-order language to be about. We've understood the widely acknowledged semantic requirement that this collection be well-specified as requiring that a domain be a non-empty extension of a collectivizing property, i.e., a sortal property that is not indefinitely extensible. Where Ψ(v) is a wff with just variable v free, ∀vΨ(v) and ∃vΨ(v) may be understood relative to an interpretation as *every (some) object in the domain that the language is about satisfies the constituent formula*. Since the domain indicator provides a reading of the quantifiers under a particular interpretation, the above may be replaced with *every (some) S satisfies the constituent formula*, where the domain indicator 'S' designates a well-defined property, e.g., '∀v' maybe read as every McKeon, every natural number, every tooth brush in Connecticut, etc.

To appreciate the import of the requirement that a domain indicator designate a collectivizing sortal property, note that it is far from obvious by the lights of the requirement that there exists a domain containing all objects since it is far from obvious that every object can be collected under a sortal property or a disjunction of them. This is noteworthy since many logicians, including Tarski, maintain that there is such a domain (e.g., Pospesel (1976) pp.50-51, J. Martin (1987), p.171, Barwise and Etchemendy (2002) pp.230-231, p.236, and Goldfarb (2003) pp. 119-121). Let's consider Tarski's argument for this view ((1941), pp.4-9, 72-74), and pinpoint the potential trouble with it. Our reconstruction of Tarski's argument runs as follows.

(1) To any sentential function with one free variable there corresponds the class of all objects satisfying the function.
(2) <u>There are sentential functions that are satisfied by every object.</u>
∴ (3) There is a class V consisting of all objects.
(4) In order to express in a first-order language that a sentential function with one free variable is satisfied by every object we must prefix a quantifier to the sentential function that binds the free variable

wherever it occurs and interpret the quantifier as ranging over V.

(5) That a sentential function with one free variable is satisfied by every object is expressible in a first-order language.

∴ (6) The class V is a domain for a quantifier prefixed to a sentential function satisfied by every object.

In order to explain the argument, consider the following true ordinary-language sentences.

(a) Paige McKeon is a Michigan resident.
(b) Paige McKeon is identical with Paige McKeon.

Replacing ‘Paige McKeon’ in both sentences with a variable, say, ‘x’, we generate what Tarski calls sentential functions.

(a') x is a Michigan resident.
(b') x is identical with x.

The non-variable terms have their ordinary meanings. Drawing on the informal characterization of satisfaction in Chapter 3, an object satisfies one of these sentential functions just in case replacing the variable wherever it occurs with a name (possibly an arbitrarily assigned name) for the object results in a true sentence. This notion of satisfaction requires that every object is nameable. Paige McKeon satisfies (a'). Barak Obama does not. Unlike (a'), (b') is satisfied by every object. Therefore, the class corresponding to (b') contains every object. Tarski calls this the universal class, and refers to it with ‘V’.

Let Ψ(v) be a sentential function with only variable v free. We may conceive of the class of things that satisfy Ψ(v) as the totality of objects that have the property *satisfies Ψ(v)*. If the class that corresponds with Ψ(v) is V, then the property of *satisfying Ψ(v)* is a logical property, and the statement that every object satisfies Ψ(v) is a statement of a logical law. Since V is the class of objects that satisfy (b'), *satisfies x is identical with x* (i.e., *being self-identical*) is a logical property of objects. That every individual x satisfies *x is identical with x* is a logical law and its representation in a first-order language (i.e., its first-order representation) would be ‘∀x(x=x)’, reading ‘∀x’ as ‘for all *objects* x’. Arguably, another candidate for a logical property is the property of *satisfying x is human or it is not true that x is human*. The first-order representation of the corresponding logical law would be ‘∀x(*x is human or it is not true that x is human*)’.

Class V is the totality of objects that logical properties are true of. The logical characterization of an object is an object that is an element of V, i.e., an object that a logical property is true of. Premise (4) tells us that the eligibility of V as a domain is essential to express in a first-order language that the relevant sentential functions are satisfied by every object. To express in a first-order language that every object satisfies Ψ(v), we must interpret '∀v' in '∀vΨ(v)' as 'every object' and regard the quantifier as ranging over V. This reflects the fact that we cannot represent the universality of the logical laws such as '∀x(x=x)' unless V is the domain over which 'x' ranges.

There is something to the idea that logical laws are fully universal. Intuitively, logical laws are true of any and all individuals. This partly motivates Quine's view that the variables are best construed as ranging over *all objects* and hence that the universal quantifier is best read 'for all objects x' and the existential quantifier as 'there exists at least one object x'. The challenge to these readings of the quantifiers is the lack of an acceptable property, or collection of them, to serve as a domain indicator. We use remarks from Martin aimed at Quine to roughly characterize the problem.

> What meaning can be given to the phrase "all objects"?] Quine, of course, prefers physical objects to others…set theorists demand another domain…but theologians need another and literary critics still another…Nothing whatsoever is gained by insisting upon the reading 'for all *objects* x'. In fact quite the contrary, a great deal is lost. What are to include as objects? Angels, dreams, dispositions, works of art, values, etc.? Russell's class of all classes that are not members of themselves? Sentences or inscriptions of such which say of themselves that they are not true? Clearly there must be some restrictions here to avoid inconsistency, on the one hand, and to gain a pittance of clarity on the other. Hence the phrase 'for all objects x' really cannot mean at all what Quine apparently wishes it to. Suitable limitations must be imposed by restricting the variables to range over some fixed domain. (1962, p. 527)

Against Tarski, it is implausible to regard V as a domain, because it is implausible to regard it as a collection whose members are just the objects which possess a sortal property (disjunctive or otherwise). There is reason to think that logical properties (as well as, *being an object*), conceived as above, are non-sortals that are non-collectivizing. This makes premise (4) problematic. Furthermore, premises (1) and (2) presuppose the truth of the conclusion (3) (how are we to read 'Every object' in (3)?), making the sub-argument circular.

Are logical properties sortals? It is far from obvious that they are associated with principles of identity and individuation. For example, in order to know what makes sentences of the form *x is the same self-identical object as y* true, one needs first to know what kind of objects *x* and *y* are, and understanding 'self-identical' does not require that knowledge (See Wright (1999) pp.314-315 for a similar consideration against treating *being self-identical* as a sortal). As Martin suggests in his criticism of Quine, the required sortal specification of "object" will reflect a particular metaphysical view of what kinds of objects there are. The long-running debate over what kinds of objects there are is rather robust, and it is unlikely that a "pittance of clarity" regarding the types of objects in a universal domain is reachable in the near future if at all. If asked to count the objects on my desk that are self-identical (or the objects on my desk that satisfy *x is human or it is not true that x is human*), I wouldn't know how to begin, because there is no principle of individuation associated with *being self-identical*. For example, every part of what is self-identical is self-identical. This provides yet another consideration against thinking that *being self-identical* is a sortal.

What we are calling logical properties, (e.g., *satisfying x is human or it is not true that x is human*, *satisfying x=x*) do not seem to be indefinitely extensible properties insofar as they are not clearly associated with a principle of extendibility. However, the universal application of logical properties (they true of any and all objects) entails that they are true of every object that possesses an indefinitely extensible property (e.g., every ordinal number is self-identical, every domain satisfies *x is human or it is not true that x is human*). Consequently, logical properties may not be indefinitely extensible, but they are nevertheless non-collectivizing, and, therefore, do not determine a domain. To illustrate, suppose that the objects that *being self-identical* is true of constitutes a domain D. Since it is a logical property, every domain is self-identical and, therefore, an element of D. Employing a version of the subdomain principle, we may derive the falsehood that there is a domain that consists of every domain. Since there is no such D, *being self-identical* is not collectivizing. A similar type of argument can be run for any logical property.

A property P can be non-collectivizing in one of two ways: (i) P itself may be indefinitely extensible, or (ii) P is not indefinitely extensible but for some indefinitely extensible property P′, all P′s are Ps. Scenario (ii) characterizes the relevant situation not only with respect to logical properties, but also with respect to the complement properties of sortals with a finite extensions (e.g., *not being a McKeon*, *not being a unicorn*, etc.). Any property P that universally holds for objects will hold for every instance of an indefinitely extensible, sortal property, thereby inheriting the non-collectivizing feature of the later.

One attempt to block the above argument is to deny the status of "objecthood" to what indefinitely properties are true of. This isn't too promising, since there are indefinitely extensible sortal properties (e.g., *being a domain* and *being an ordinal number*) and on the model-theoretic account of the general concept of "object" what sortal properties are true of are objects. Furthermore, ordinals and domains are, intuitively, self-identical and so should qualify as objects by Tarski's lights. This is in tension with Tarski's characterization of V as the extension of the general concept of "individual" (i.e., "object"), and not the class of all *possible things* (e.g., classes, classes of classes, etc.) (1941), p. 73). Other proponents of a universal domain of objects also rule out classes as elements (e.g., Pospesel (1976), p. 51). Again, these moves are questionable, since *being a class* is a sortal property whether we understand "class" as another name for a set (aren't logical properties true of a class of classes?) or understand it as referring to both sets and non-set collections, i.e., *proper classes*. If we treat *being an ordinal* and *being a domain* as collectivizing properties by regarding the collections of their instances as proper classes (so, for example, the collection of all domains is not itself a domain), then it is hard to see why such collections are excluded from V since logical properties are true of proper classes. Certainly, in order to respect the universality of logical laws, we should think that they hold for all *possible things*, where 'things' means *objects*. In short, V, specified as the totality of what logical properties are true of, cannot serve as a domain, because logical properties are what we are calling hazy.

It is worth summarizing why classical quantification over indefinitely extensible totalities is problematic. Shapiro and Wright ((2006), pp.296-297) put it as follows.

> Think of quantification as a many-to-one function that yields a truth-value when given a range of instances as argument. If the elements which the instances concern are indefinitely extensible, then no application of the function can embrace them all—for any collection of the instances to which it is applied will immediately be at the service of the definition of a new instance, so far unreckoned with.
>
> The crucial thought is thus that a function requires a *stable* range of arguments if it is to take a determinate value. The operation of classical quantification on indefi-

> nitely extensible totalities is frustrated because any attempt to specify them subserves the construction of a new case, potentially generating a new value. The reason why quantification, classically conceived, requires a domain to operate over is just that.

Let's formulate the insight here in terms of the semantics for our language M. A variable assignment is a function *g* from a set of variables (*g*'s domain) to a set D of objects (*g*'s range). With respect to a language L and domain D, a variable assignment assigns elements of D to L's variables. Recall that the truth clauses (VI) and (VII) for the quantifiers require that each element of the domain is a possible value of a variable, i.e., they require that for any variable *v*, assignment *g*, and element *d* of D, either g(v)=d or *g′*(v)=d for some extension *g′* of *g*. The range of a variable assignment cannot be an indefinitely extensible totality, because it must be *stable* if quantification is to take a determinate truth-value. Any attempt to determine a range R of a variable assignment by specifying a non-collectivizing sortal generates a new object outside R potentially changing the truth-value of a quantification where R to be expanded to include it. Not all quantifications are potentially indeterminate. Purely existential quantifications (quantifications lacking a universal quantifier) are persistent through expansion. That is, whenever a purely existential quantification is true under an interpretation relative to a domain D, it remains true under every expansion of D. Also, Dummett ((1981), p.531) remarks that sentences such as ∀x∃y(x=y) and ∀x(Male→~Female(x)) are determinately true even if the range of the variables is an indefinitely extensible totality, since there is no potential for their truth-value relative to a domain D to change due to expansions of D.

The fact that being a domain is an indefinitely extensible property poses a serious challenge to the model-theoretic characterization of logical consequence. We have defined (in the meta-language) the model-theoretic consequence relation for our language M: X is a model-theoretic consequence of K iff (1) for every M-structure U, if the K-sentences are true relative to U, then X is true relative to U. In other words: K ⊨ X iff (2) there is no domain D and interpretation I of M such that the K-sentences are true and X is false in D under I. We take this to capture the idea that if X is a logical consequence of K, then (3) for every domain-like collection C of objects, if M's variables range over C, its individual constants refer to objects in C and its predicates designate subsets of C, X is true if the K-sentences are true. Note that in (1)-(3), we quantify over structures, domains and interpretations, and hence ostensibly we treat these things as objects, and require that the collection of all domains is itself a domain. But, as discussed above, this requirement is problematic.

As summarized above, the classical model-theoretic semantics for the quantifiers is threatened by the fact that the property of being a domain (as well as the properties of being an interpretation and being a structure) is non-collectivizing. The viability of quantifying over indefinitely extensible totalities is a substantive issue (see the anthology edited by Rayo and Uzquiano (2006)). Here we briefly consider two promising strategies for deflecting this challenge which propose that we abandon the classical model-theoretic treatment of quantification in understanding the generality in the model-theoretic characterization of logical consequence.

The first strategy is to justify quantification over indefinitely extensible totalities by arguing against the requirement that a domain indicator must designate a collectivizing property. Cartwright calls the requirement that in order to quantify over objects those objects must form a set-like collection the All-in-One Principle ((1994), p.7). The All-in-One principle reflects the requirement that what first-order properties range over is the extension of a collectivizing property. Cartwright rejects the All-in-One principle in favor of allowing quantification over objects even though those objects do not constitute one set-like object (i.e., even though there is no one thing of which they are members). Admittedly, Cartwright's view does have an easy-to-see plausibility: it is seemingly counterintuitive to require that my assertion that, say, some cookies in the jar are chocolate chip, commits me to the existence of a set-like collection of cookies in addition to the cookies. According to Cartwright, the All-in-One principle, presupposed by model-theoretic semantics, is merely an artifact of the model-theoretic characterization of logical consequence, and as such does not justify thinking that quantification over indefinitely extensible totalities is illegitimate. Drawing on an argument from Kreisel, Cartwright notes in effect that a first-order sentence X is a logical consequence of a set K of first-order sentences iff there is no totality D that is the extension of a collectivizing, sortal property and interpretation I according to which all K-sentences are true and X is false relative to D. This suggests that we can reject the All-in-One principle while believing that the model-theoretic characterization of logical consequence is extensionally adequate.

The second strategy proposes that we regard the generality in the model-theoretic characterization of logical consequence as schematic and not quantificational, where the former is not reducible to the later. Our explication of schematic generality follows Lavine (1996). Let ⊨ be defined for some first-order language L. The letters in the meta-linguistic schematic generalizations, {P, ~P} ⊨ Q and ⊨ P V ~P, are schematic variables. These schemes can be used to assert that *any* of their instances is true, where *any* is to be understood as indicating schematic and not quantificational generality. Unlike the referential variables of our interpreted language M, schematic variables do not range over a domain of elements. With reference to the above two schemes, instances are truth-bearers (e.g., sentences). The presence of free schematic letters in the schemes prevents them from being truth bearers. To "assert" the two schemes commits one to the truth of all their instances. The schemes are open-ended in the sense that what counts as an acceptable sentential substitution instance is open-ended, and expands as L expands to include new terminology. This is why Lavine refers to them as full schemes ((1996), p. 136).

> "To understand my analysis, it is important to recall that I take the generality of schemes, including full schemes to be *sui generis*, not to be regarded as defined using quantifiers and that full schemes are stronger than their universally quantified counterparts since full schemes express truths encompassing objects that need not be in the current universe of discourse."

Employing the idea of full schemes, the meta-linguistic characterization of logical consequence for a language L may be put as follows.

For any K and any X, K ⊨ X iff any structure <D,I> that is a model of the K-sentences is a model of X.

K and X are schematic variables whose instances are any set of L-sentences and any L-sentence, respectively. D and I are schematic variables of the meta-language whose instances are any domain indicator and any expression for an interpretation function, respectively. Since full-schematic generality is basic and not dependent on any form of quantificational generality, the generalization does not call upon a domain of discourse consisting of sets of L-sentences, L-sentences, domain indicators and interpretation functions. Therefore, sentences from expansions of L, as well as domain indicators and expressions that designate interpretations in expansions of the metalanguage are potential instances of the variables.

Interpretations

Our introduction to structures illustrates how model-theoretic semantics allows for three degrees of fixity or variation in the semantic values of a formal language L's alphabet (Peters and Westerstahl (2006), p. 41-42).

(1) Certain semantic values (e.g., of logical constants) are completely fixed given by the semantic rules for truth in arbitrary structures.
(2) Other semantic values (e.g., of individual constants and predicates) are relatively fixed, given by a particular model's domain and interpretation function.
(3) The remaining ones are hardly fixed at all, given only by variable assignments to free variables.

Variations in the semantic values of individual constants, predicates, and variables represent possible uses of the relevant language. By virtue of depicting possible uses for language, structures are representations of logical possibilities. That a set K of sentences are true and sentence X is false relative to a structure shows that X is not a logical consequence of the K-sentences, because the possible use for the relevant language depicted by the structure represents the logical possibility of the K-sentences being true and X being false. By our lights, it is logically possible that every sentence in a set K of sentences be true iff there is a possible use of the relevant language according to which they are all true. In what follows, we clarify the notion of a possible use for language that the second and third degrees of variability in semantic value represent, and highlight the role that interpretations play in representing such uses.

We've seen that by varying the semantic values of some of M's alphabet, we demonstrate that one sentence is not a model-theoretic consequence of others. We produce a simple illustration to refer to in what follows. Consider the following two sentences.

(a) Parent(beth, paige)
(b) OlderThan(beth, paige)

Both are true relative to U^M, i.e., the intended structure for M. Here 'Parent' and 'OlderThan' have their ordinary designations defined over the McKeons. The term 'Beth' refers to Beth and 'Paige' refers to Beth's daughter. Sentence (b) is not a logical consequence of (a). To show this, consider an M-structure U' which we partially describe as follows. The domain D = {1, 2}, 'OlderThan' designates the >-relation, and 'Parent' the <-relation. Let 'beth' refer to 1 and 'paige' to 2. What makes the fact that 1 is less than and not greater than 2 relevant to establishing that (b) is not a logical consequence of (a) is the semantic possibility of 'Parent' and 'OlderThan' having the above extraordinary designations, and the semantic possibility of the individual constants referring to 1 and 2, i.e., referring to objects other than what they refer to in U^M. We demonstrate that (b) is not a logical consequence of (a) by appealing to the (semantic) possibility that constants and predicates designate things other than what they designate under the intended interpretation. In short, that there are possible uses or meanings for expressions such as 'beth', 'paige', 'Parent' and 'OlderThan' according to which (a) would be true and (b) false reflects that the semantic value of an expression is a contingent (or, equivalently, accidental) feature of it. More generally, the second and third degrees of variability in the semantic values of a formal language L's alphabet allowed for in model-theoretic semantics reflects the view that the connection between individual constants, predicates, and variables on the one hand and their semantic values on the other is contingent. This view is orthogonal to semantic essentialism, which is the view that words have their semantic values essentially.

We have seen that sentences (a) and (b) are true relative to U^M whereas (a) is true and (b) is false relative to U′ depends partly on the fact that $I_{U'}$ (beth)=2 and $I_U{}^{\mathbf{M}}$ (beth)= Beth. We understand this fact as one word, 'beth', being interpreted in two different ways, each interpretation reflecting a possible use for the word. However, according to semantic essentialism, $I_{U'}$ (beth) and $I_U{}^{\mathbf{M}}$ (beth) are two different words. On this view, constants are individuated by the semantic values assigned to them under an interpretation.

In order to reflect semantic essentialism, we may take M to be (syntactic) template for a language. A U^i-structure for M generates a language M^i, an instance of the template M. The class of structures used to interpret M generates a family of languages, each an instance of M.

M + U′ = M′
M + U″ = M″
M + U‴ = M⁗
and so on.

What distinguishes a language is the structure that, in conjunction with M, generates it. Let X be a wff from template M and K a set of M-wffs. On this approach, for a given language M^i, a sentence X^i is a logical consequence of a set K^i of sentences iff there is no language M^j such that the K^j-sentences are true and X^j is false, where X^i and X^j are instances of X and the K^i-sentences and K^j -sentences are instances of the K-wffs. Hopefully, enough has been said about semantic essentialism to contrast it with the semantic approach to formal languages adopted here.

In Chapter 3, we said that a formal language such as M can be identified with the set of its wffs, and is specified by giving its alphabet of symbols and its formation rules. The formation rules determine which strings constructible from the alphabet are wffs. Since both the alphabet and formation rules can be completely specified without any reference to interpretation, M, like all formal languages, can be characterized syntactically, i.e., can be defined without essential regard to its interpretation. M's vocabulary items are essentially syntactic entities, and, therefore, retain their identities through various interpretations of them. An expression such as 'beth' retains its identity across interpretations of it; 'beth' is a syntactic entity.

The variability in the semantic values of M's terms and predicates reflects that these expressions have possible meanings. Possible meanings or uses for terms and predicates aren't that strange. The connection between a word and its ordinary use is, after all, a matter of convention, e.g., the word 'married' didn't have to mean *married*. Also, many words have more than one use that is possible relative to the conventions of English. For example, indexicals like 'you' and 'here' change their referent with occasion of use; there are many people named 'Bill'; and 'married' can be used to mean *was married to* or *performed the marriage of.*

If there were no restrictions on possible meanings or uses, then no sentence would be a logical consequence of others, which is, of course, undesirable. We do not want things arranged so that '∀x Female(x)' fails to be a logical consequence of '∀y Female(y)' because '∀' in the first sentence can be used to range over a domain where 'Female' is true of everything, and '∀' in the second sentence can be used to range over a domain where it is true of nothing. And we shouldn't think that 'Parent(beth, paige)' is not a logical consequence of 'Parent(beth, paige) & OlderThan(beth,paige)' because '&' could be used to mean '∨'.

In specifying a use for a first-order language L, we specify both a domain of discourse and an interpretation in the domain of the non-logical elements of L's alphabet. Also, the domain of discourse serves as a parameter for a possible use of the sundry variables in L; a possible use for the L variables is given by an assignment which makes the variables range over the domain. So, for example, there is no possible use for ‘x’ and ‘y’ according to which ‘∀x Female(x)’ is false and ‘∀y Female(y)’ true because, in part, a possible use for all variables from language M is fixed in terms of the one collection selected as the domain of discourse for M. Relatedly, the domain of discourse for a language L is a determinant of a possible use of L’s names and predicates because the extension of these alphabet-expressions on a given use of them must come from the elements of the domain. In general, the domain of discourse serves to coordinate the possible uses of a language L’s non-logical elements as part of the global interpretation L. For example, although a possible use for ‘Female’ is to refer to any set, its use must be coordinated in the right way with a use of ‘∃’ in order for the sentence ‘∃x Female(x)’ to receive a truth-value on a given use of it, i.e., ‘Female’ must be used to pick out a subset of the domain as this is fixed by a given use of ‘∃’. Possible uses for the non-logical elements of a language’s alphabet (i.e., variables, individual constants, predicates, possibly functional expressions) are to be constrained only by the type of expression they are (e.g., predicates must be used to pick out a subset of the domain, individual constants must refer to elements of the domain, variables must range over the domain, etc.).

In Chapter 2 we said that the logical form (or logical structure) of a sentence is determined by the logical constants that appear, if any, and the pattern of the remaining non-logical elements. The logical form of a sentence determines how the truth or falsity of the sentence depends on a use of its non-logical components as this is coordinated by a domain of discourse. For example, taking ‘&’ to be the logical constant in ‘Parent(beth, paige) & OlderThan(beth,paige)’ determines that it is only true on possible uses for the components of ‘Parent(beth, paige)’ and ‘OlderThan(beth,paige)’ according to which they both are true. A possible use for a sentence will be any coordinated use of its non-logical components consistent with its logical form. Hence, there is no possible use for ‘Parent(beth, paige) & OlderThan(beth,paige)’ which treats ‘&’ as ‘∨’, because such a use ignores the sentence’s logical form and allows uses for ‘Parent(beth, paige)’ and ‘OlderThan(beth,paige)’ according to which ‘Parent(beth, paige) & OlderThan(beth,paige)’ is true when ‘Parent(beth, paige)’ is false. But this is inconsistent with the fact that ‘Parent(beth, paige) & OlderThan(beth,paige)’ has a conjunctive logical form.

An account of possible meaning

Here an account of an alternative possible meaning of a term or predicate expression is offered from which we derive the notion of a possible use for a formal language. On way of understanding possible meaning calls upon ways the world could have turned out. For example, one could think that in order for, say, 'carrots' to have designated apples the brain states of people who had a causal role in the way the word is actually used would have to have been different. We aim to avoid possible worlds in constructing an account that reflects the above considerations regarding constraints on the possible meanings of terms and predicates.

When the author of an artificial language assigns an expression to a syntactic category, the author assigns it a certain function/role in the construction of formulae. Contra Wittgenstein of the *Tractatus*, the *metaphysical essence* of an expression, i.e., it being a certain orthographic/phonographic type, does not determine its assigned syntactic category. For example, nothing in the nature of 'OlderThan' prohibits it from being used as a constant to pick out an individual. Since an expression's syntactic category is conventionally determined, we shall refer to it as the expression's *conventional essence*. An interpretation of a language relates the language to the world, by assigning semantic values (i.e., things in the world) to the expressions of the language's alphabet. Therefore, an account of how such expressions can be related to the world constrains the interpretations that the language can have. Such an account assigns to each syntactic category C a semantic type that determines a range of semantic values that any element of C has the potential to possess. E.g., each term has the potential to refer to a particular individual, each n-place predicate the potential to designate a set of ordered n-tuples of individuals.

Syntactic Category	*Semantic Type*
individual constant	individual
1-place predicate	set of individuals
2-place predicate	set of ordered pairs of individuals
3-place predicate	set of ordered triples of individuals
n-place predicate	set of ordered n-tuples of individuals

The semantic potential of an expression e of syntactic category *C* is the range of semantic values any one of which e may possess by virtue of being a C. An expression e's potential to possess a particular semantic value v is explained by e being in the syntactic category C that corresponds to the semantic type of which v is an instance. We regard an alternative possible use/meaning of an expression as the expression's potential—relative to its assignment in a syntactic category—to have that use/possess that meaning. We understand an expression's potential use (i.e., its semantic potential) as follows. An expression e has the potential to possess semantic value v iff it is not essential (in the conventional sense) to e that e not possess v. It is essential to 'OlderThan' *qua* predicate that it not refer to Shannon, it is essential to 'shannon' *qua* individual constant that it not correspond to what is true of individuals. Possible uses of variables, individual constants, and predicates are to be constrained only by the syntactic type of expression they are (e.g., predicates must be used to pick out a subset of the domain, individual constants must refer to elements of the domain, variables must range over the domain, etc.).

The actual semantic value of an expression is a contingent (or accidental) feature of it, because it has the potential to possess a different semantic value of the same semantic type. Following Fine (1994) and Gorman (2005), the essence/accident distinction is not understood here in modal terms (according to which a property of an object is essential to it if it cannot exist without it; a property of an object is contingent if it can exist without it.) Membership in a syntactic category is an essential characteristic of an expression *e* (its conventional essence), because it is not explained by any other characteristics of *e*. (See Gorman's account of the essence/accident distinction (2005) pp. 282-285, which we utilize in the following paragraph.)

A property of an expression that explains its semantic potential without itself being explained by any other characteristic of the expression is an essential property of the expression. An essential property of an expression is directly definitive of the expression; it is not had in virtue of being a consequence of other properties that characterize the expression. To say that one thing explains another is not necessarily to say that it is a complete explanation of it. Also, to say that an essential property of x is not explained by any other characteristic of x is not to say that it is inexplicable. The explanation relation that I am highlighting is: an expression being a member of a syntactic category explains its semantic potential. To speak of explanation in this sense is to use the word in its ontic sense, not its epistemological sense. To say that an expression's syntactic category explains its semantic potential (it being in a certain syntactic category is why it has the semantic potential it has) is to make a claim about a real relation/connection between the expression's syntactic category and its meaning (potential or actual). Whether or not anyone ever uses the expression's syntactic category to explain its semantic potential, the former does indeed explain the later.

A possible use of a language L is L's potential to be about a state of affairs S by virtue of the potential of L's terms and predicates to have as semantic values the constituents of S. The notion of a possible use of variables and therewith quantifiers can be used to clarify what state of affairs is represented by a domain that is a restriction of all that there is. By allowing truth-value assignments to quantificational sentences that are relative to any non-empty subset of the world's first-order individuals, model-theoretic semantics represents the fact that there are different ways of assigning portions (i.e., fragments) of the actual world as meanings for quantifiers. For example, we needn't think that the only way for the sentence, 'Everybody is large' to be true while keeping the meaning of 'large' fixed is for it to be interpreted in a merely possible world whose residents are all mesomorphs. Rather we imagine a context in which an utterance of the sentence would make it actually true. For example, while looking at a picture of NFL football players a person comments to another, 'Everybody is large'. The context here determines that the quantifier 'everybody' refers to those pictured, not to all human beings. Clearly, 'Everybody is large' has at least as many uses as there are domains for 'Everybody' (for a contrary view see Bach (2000)).

The idea of a possible use for an ordinary language quantifier highlights the fact that (assuming an existent infinity) for each positive integer n, a possible use of the variables and associated quantifiers in a quantificational sentence is to range over just n-1 of the things that, in fact, exist (as applied to classical logic, n>1). By assigning a truth-value to a quantificational sentence relative to a subset of worldly objects, we are not considering different ways earth, air, fire, and water could have been distributed. Rather, we represent one way the variables of our language can range over what, in fact, exists. This is the philosophical rationale for the model-theoretic use of any non-empty set as a domain. There is no appeal to possible worlds that portray the world with finitely many individuals. Different quantifier ranges do not represent the world with different numbers of individuals; they represent possible uses for the quantifiers according to which they range over a subset of what, in fact, exists. Understanding 'truth simpliciter' as 'as truth in the actual world' as opposed to 'truth in a merely possible world', *truth relative to a domain that is a proper subset of what exists* is truth simpliciter.

This doesn't seriously threaten the status of the universal quantifier as a logical constant, because there are two elements to the meaning of quantifier only one of which is completely fixed in terms of the semantic rules for truth in arbitrary structures. Consider the following.

(1) $\forall x\ \psi(x)$
(i) ***For all*** *individuals* x, $\psi(x)$
(2) $\exists x\ \psi(x)$
(ii) ***For some*** *individuals* x, $\psi(x)$

The meaning of the quantifier symbols is italicized: '$\forall$' and '$\exists$' always mean 'for all individuals' and 'for some individuals', respectively. What is in bold is the logical element of quantifier meaning, and is fixed from one structure to another. However, what counts as an individual depends on how we are using the language. A domain of quantification represents a possible use for 'individual'. A possible use for the quantifier symbols of a language L is determined by a particular domain of discourse for L. As L may have different domains of discourse (i.e., may have different applications), there are different uses for the quantifier symbols of L. The fact that the quantifiers are logical constants means that the part of their meaning that is fixed is the portion of the domain they pick out (e.g., the existential quantifier refers to at least one member of the domain and the universal quantifier refers to all of it), and the part of their meaning that can be varied (i.e., the non-logical element of the quantifiers) in order to make a sentence true or false is the size of the domain. So, we may regard the existential quantifier as a logical constant and still think that a sentence like '$\exists x \exists y \sim(x=y)$' is not a logical truth, since a possible use of the variables is to range over exactly one thing, and when the sentence is used this way it is false.

Model-theoretic truth is a relativized conception of truth, e.g., '$\exists x \exists y \forall z(x \neq y\ \&\ (z=x \vee z=y))$' is true relative to a structure whose domains has only two elements, false relative to others. However, the truth of an interpreted sentence in a domain smaller than that of the collection of all individuals is nevertheless truth simpliciter, i.e., truth in the actual world. To see this clearly, the model-theoretic characterization may be portrayed as follows. Let K be a set of finitely many sentences and X a sentence. The antecedent of a corresponding conditional is the conjunction of the K-sentences, while its consequent is X. Generate a new sentence S by uniformly replacing each individual constant occurring in this conditional by a first-order variable, each one-place predicate 'Fv' by '$v \in F$' (i.e., 'v is a member of F'), each two-place predicate 'Rvu' by '$<v,u>\in R$' (choosing new variables each time) and so on. Then restrict all variables to a new second-order variable 'X', which ranges over the universe of sets. To do this, make sentence S the consequent of a new conditional in whose antecedent put '$v\in X$' for each individual variable 'v', and '$V\subseteq X^n$' (i.e., 'V is a set of ordered n-tuples of members of X') for each n-adic second-order variable 'V', and then universally close. The general form of the resulting sentence will be as follows.

$\forall X,\forall v_1, \ldots,\forall v_n,\forall V_1, \ldots\forall V_n[(X\neq\varnothing\rightarrow(v_1\in X, \&\ldots\& v_n\in X \& V_1\in X \& \ldots \& V_n\in X))\rightarrow S]$

The variables v_i and V_i are the free first-order and second-order variables occurring in S. Then $K \models X$ iff the corresponding set-theoretic sentence is true simpliciter. For example, '$\forall xFx$' is not a model-theoretic consequence of '$\exists xFx$' because

'$\exists X\exists x\exists y\exists F(X\neq\varnothing \;\&\; x\in X \;\&\; y\in X \;\&\; x\neq y \;\&\; F\subseteq X \& \; x\in F \;\&\; y\notin F)$'

is true where the quantifiers range over all the entities of the appropriate type, i.e., 'x' and 'y' range over the totality of first-order elements, 'X' and 'F' range over the totality of sets of first-order elements. Since no non-logical terminology appears in (i) $\exists x\exists y\; x\neq y$ and (ii) $\exists x\exists y\exists z\; (x\neq y \;\&\; y\neq z \;\&\; x\neq z)$, the model-theoretic explanation why (ii) is not a logical consequence of (i) turns on the fact that that there exists a set with just n objects, i.e., turns on the fact that

'$\exists X\exists x\exists y\forall z\; [(X\neq\varnothing \;\&\; (x\in X \;\&\; y\in X \;\&\; x\neq y)) \;\&\; (z\in X \rightarrow(z=x \vee z=y))]$'

is true.

The Adequacy of the Model-Theoretic Characterization of Logical Consequence

Our aim in this chapter has been to characterize logical consequence for language M in terms of the model-theoretic consequence relation. If successful, the model-theoretic characterization is extensionally equivalent with the common concept of logical consequence as applied to M. In order to set up a framework for evaluation of the extensional adequacy of the model-theoretic characterization, we review the key moves in our Tarskian analysis of logical consequence for first-order languages such as M. In what follows, K is an arbitrary set of sentences from a language L, and X is any sentence from L.

First, we record Tarski's observation of the salient features of the common concept.

X is a logical consequence of K iff *(a) it is not logically possible for all the sentences in K to be true and X false, (b) this is due to the forms of the sentences, and (c) this is knowable a priori.*

The second step in Tarski's analysis is to embody (a)-(c) in terms of the notion of a possible interpretation of the non-logical terminology in sentences. Substituting for what is italicized, we get Tarski's precisifying definition of logical consequence.

X is a logical consequence of K iff *there is no possible interpretation of the non-logical components of the language according to which all the sentences in K are true and X is false.*

The third step of our Tarskian analysis of logical consequence is to employ the modern, technical notion of *a structure* or *model*, a descendent of Tarski's (1936) article, to capture the idea of a *possible interpretation*. That is, we understand *there is no possible interpretation of the non-logical terminology of the language according to which all of the sentences in K are true and X is false* in terms of: *Every model of K is a model of X, i.e., K* $\models$ *X.*

To elaborate, as reflected in the second step, the analysis turns on a selection of terms as logical constants. This is represented model-theoretically by allowing the interpretation of the non-logical terminology to change from one structure to another, and by making the interpretation of the logical constants invariant across the class of structures. Then, relative to a set of terms treated as logical, the Tarskian, model-theoretic analysis is committed to the following.

X is a logical consequence of K iff K $\models$X

Moving beyond Tarski, we take a possible interpretation of the non-logical terminology of a language to be a possible use of it, and understand what's italicized in Tarski's precisifying definition as *there is no possible use of the non-logical components of the language according to which all of the sentences in K are true and X is false*. A structure for a language represents a possible use of the language.

From the above analysis of logical consequence, we derive a case for the extensional adequacy of the model-theoretic characterization.

(1) X is a logical consequence of K iff it is not logically possible for all the sentences in K to be true and X false, this is due to the forms of the sentences, and this is knowable *a priori*.

(2) It is not logically possible for all the sentences in K to be true and X false, this is due to the forms of the sentences, and this is knowable *a priori* iff there is no possible use of the non-logical components of the language according to which all of the sentences in K are true and X is false.

(3) There is no possible use of the non-logical components of the language according to which all of the sentences in K are true and X is false iff K $\models$X.

Therefore, X is a logical consequence of K iff K $\models$X.

We regard the conclusion as a statement of the extensional adequacy of the model-theoretic characterization of logical consequence. Premises (1) and (2) are analyses of the left sides of the iffs, and are, therefore, necessarily true if true. Like the conclusion, premise (3) is an extensional equivalence. In what follows, we consider challenges to (1)-(3), and sketch responses. Then we consider what the argument suggests about the status of the model-theoretic characterization of logical consequence. It is far from obvious that it is an analysis of the concept of logical consequence. The above argument suggests that its use to fix what follows logically from what presupposes such an analysis as represented in (1) and (2) from which it follows that *X is a logical consequence of K iff there is no possible use of the non-logical components of the language according to which all of the sentences in K are true and X is false*. This has the status of a theoretical definition. One question raised here is: how exactly does the model-theoretic characterization engender understanding of logical consequence? After responding to this question, we end the chapter with discussion of what qualifies as a logical constant.

Premise (1): Is Tarski's identification of the salient features of the common concept of logical consequence correct?

Philosophers and logicians differ over what the features of the common concept are. Some offer accounts of the logical consequence relation according to which it is not *a priori* (e.g., see Koslow (1999), Sher (1991) and see Hanson (1997) for criticism of Sher) or deny that it even need be strongly necessary (Smiley 1995, 2000, section 6). Here we illustrate with a quick example.

Given that we know that a McKeon only admires those who are older (i.e., we know that (a) ∀x∀y(Admires(x,y)→OlderThan(y,x))), wouldn't we take (b) to be a logical consequences of (c)?

(b) Admires (paige, kelly)
(c) OlderThan (kelly, paige)

A Tarskian response is that (c) is not a consequence of (b) alone, but of (a) plus (b). So in thinking that (c) follows from (b), one assumes (a). A counter suggestion is to say that (c) is a logical consequence of (b) for if (b) is true, then *necessarily-relative-to-the-truth-of-(a)* (c) is true. The modal notion here is a weakened sense of necessity: *necessity relative to the truth of a collection of sentences*, which in this case is composed of (a). Since (a) is not *a priori*, neither is the consequence relation between (b) and (c). The motive here seems to be that this conception of modality is inherent in the notion of logical consequence that drives deductive inference in science, law, and other fields outside of the logic classroom. This supposes that a theory of logical consequence must not only account for the features of the intuitive concept of logical consequence, but also reflect the intuitively correct deductive inferences. After all, the logical consequence relation is the foundation of deductive inference: it is not correct to deductively infer B from A unless B is a logical consequence of A. Referring to our example, in a conversation where (a) is a truth that is understood and accepted by the conversants, the inference from (b) to (c) seems legit. Hence, this should be supported by an accompanying concept of logical consequence.

This idea of construing the common concept of logical consequence in part by the lights of basic intuitions about correct inferences is reflected in the Relevance logician's objection to the Tarskian account. The Relevance logician claims that X is not a logical consequence of K unless K is relevant to X. For example, consider the following pairs of sentences.

(d) (Female(evan) & ~Female(evan))	(d) Admires(kelly, paige)
(e) Admires(kelly, shannon)	(e) (Female(evan) v ~Female(evan))

In the first pair, (d) is logically false, and in the second (e) is a logical truth. Hence it isn't logically possible for (d) to be true and (e) false. Since this seems to be formally determined and *a priori*, for each pair (e) is a logical consequence of (d) according to Tarski. Against this Anderson and Belnap write, "the fancy that relevance is irrelevant to validity [i.e. logical consequence] strikes us as ludicrous, and we therefore make an attempt to explicate the notion of relevance of A to B" (Anderson and Belnap (1975) pp. 17-18). The typical support for the relevance conception of logical consequence draws on intuitions regarding correct inference, e.g. it is counterintuitive to think that it is correct to infer (e) from (d) in either pair for what does being a female have to do with who one admires? It seems incorrect to infer, say, that *Admires(kelly,shannon)* on the basis of *(Female(evan)&~Female(evan))*. For further discussion of the different types of relevance logic and more on the relevant philosophical issues see Haack (1978) pp. 198-203 and Read (1995) pp. 54-63. The bibliography in Haack (1996) pp. 264-265 is helpful. We say more about relevance logic below in Chapter 5.

The considerations here suggest that the deductive-theoretic characterization of logical consequence is prior to the model-theoretic characterization. If this is right, then we may justify the extensional correctness of the model-theoretic characterization (the *if* part of the conclusion of the argument for the adequacy of the model-theoretic characterization) by appealing to the correctness of a deductive system which captures the inferential properties of the logical expressions of M. The following argument illustrates this.

(f) If K $\models$ X, then K $\vdash_D$ X.
(g) If K $\vdash$ X, then X is a logical consequence set K.
So, if K $\models$ X, then X is a logical consequence of set K.

Recall from Chapter 2 that 'K $\vdash_D$ X' means that X is a deductive-consequence in a deduction system D of K. The deductive-theoretic characterization of logical consequence and its relation to the model-theoretic characterization is discussed in Chapter 5. Here we simply note that premise (f) is the completeness theorem for $\vdash_D$, and premise (g) presupposes the primacy of $\vdash_D$ over $\models$ (which Tarski rejects), and requires philosophical justification of $\vdash_D$.

One reaction to the dispute over Tarski's identification of the salient features of logical consequence is to question the presupposition of the debate and take a more pluralist approach to the common concept of logical consequence. On this line, it is not so much that the common concept of logical consequence is vague as it is ambiguous. At minimum, to say that a sentence X is a logical consequence of a set K of sentences is to say that X is true in every circumstance [i.e. logically possible situation] in which the sentences in K are true. "Different disambiguations of this notion arise from taking different extensions of the term 'circumstance' " (Restall (2002 p. 427). If we disambiguate the relevant notion of 'circumstance' by the lights of Tarski, 'Admires(kelly, paige)' is a logical consequence of '(Female(evan) & ~Female(evan))'. If we follow the Relevance logician, then not. There is no fact of the matter about whether or not the first sentence is a logical consequence of the second.

Historically, logical consequence has been understood to be knowable *a priori*, formal, and necessary. Certainly, Tarski's identification of these as the salient features of logical consequence is not idiosyncratic. A pluralist approach to the common concept of logical consequence is in sync with the idea alluded to in Chapter 2 that the concept of logical consequence that Tarski aims to make precise is one common to researchers of his time in the foundations of mathematics. The debate over the accuracy of Tarski's informal characterization of the salient features of the concept of logical consequence turns on identifying the version of this concept prevalent in the community of researchers in the foundations of mathematics. As such, the debate seems to be a historical matter to be settled by empirical investigation.

Premise (2): Does Tarski's precisifying definition of logical consequence capture the modal element in the concept of logical consequence?

We now consider criticism of premise (2) in the form of a challenge to its portrayal of the modal element in the concept of logical consequence. Premise (2) entails that it is necessarily the case that it is logically possible for all sentences in a set K to be true with a sentence X false only if there exists a possible use of the non-logical components of the K-sentences and of X according to which the former are all true and the later false. We develop a criticism of premise (2) which rejects this and which provides a rationale for a negative response to the above critical question. We shall reply in defense of (2).

Even on a pluralist approach to the common concept of logical consequence, it has to be acknowledged that the usefulness of the logical consequence relation, on any common understanding of it, lies in the fact that if a sentence X is a logical consequence of a set K of sentences, then X is guaranteed to be true if the K-sentences are true, i.e., it is guaranteed that truth is preserved from K to X. Let's call this the truth-preservation guarantee of logical consequence, for short *the truth-preservation guarantee*. The truth-preservation guarantee is a *logical* guarantee of the truth of X given the truth of the sentences in K (and thus determinable *a priori*), because the supposition that X is false and all the K-sentences are true violates the meanings (as determinants of the truth-conditional properties) of component *logical* terminology. This suggests that the source of the truth-preservation guarantee is the modal element in the concept of logical consequence. A logical consequence X of a set K of sentences is guaranteed to be true if the K-sentences are, because it is logically impossible for X to be false with all the K-sentences true.

Appealing to our account of logical possibility reflected in (2), truth is preserved from K to X, because there are no possible uses for the non-logical components of the sentences in K and of X according to which the sentences in K are true and X is false. We summarize this in the form of (i).

(i) If there is no possible use of the non-logical components of the sentences in K and of X, according to which the K-sentences are true and X is false, then X is logically guaranteed to be true if all the K-sentences are true.

It seems that in order for (i) to be true, (ii) must be true.

(ii) The truth-conditional properties of the occurrent logical expressions provide the assurance that there is no possible use of the non-logical components of K and X according to which the former are true and the latter is false.

The idea here is that the source of the truth-preservation guarantee is the truth-conditional properties of logical expressions. For now, we regard (ii) as the reason why there being no possible use of the non-logical terminology of K and X according to which all K-sentences are true and X is false logically guarantees the truth of X if the K-sentences are true. Proposition (ii) highlights that the supposition that there is a possible use of the non-logical terminology of K and X according to which K is true and X is false violates the meanings of the occurrent logical expressions as this is reflected in their truth-conditional properties. We illustrate by way of a simple example.

The sentence 'OlderThan(beth, matt) ∨ Male(evan)' is a logical consequence of 'Male(evan)'. Suppose that there is a possible use for 'Male', 'OlderThan', 'beth', 'matt', and 'evan' according to which 'Male(evan)' is true and 'OlderThan(beth, matt) ∨ Male(evan)' false. If the later, then by the semantics for ‘∨’, ‘Male(evan)’ is false contrary to our initial hypothesis. This demonstrates that a possible use for the predicates and individual constants is impermissible by the relevant truth-conditional property of ∨ as spelled out by ∨’s truth clause. ‘Male(evan)’ is not a logical consequence of ‘OlderThan(beth, matt) ∨ Male(evan)’, because a possible use of the predicates and individual constants is permissible by the truth-conditional property for ∨. Since possible uses for a language are constrained by the meanings of the logical expressions, there being no possible use of the non-logical components of K and X according to which the former are true and X is false is the source of the logical guarantee of the truth of X if all K sentences are true.

There are critics of Tarski’s precisifying definition of logical consequence who argue that it fails to account for the modal element of the concept of logical consequence, because it lacks the conceptual resources to rule out there being logically possible situations in which sentences in K are true and X is false, but no possible interpretation of the non-logical terminology of the language according to which all the sentences in K are true and X is false. Kneale (1961) is an early critic, Etchemendy (1988, 1999a, 1999b) offers a sustained and multi-faceted attack. We follow Etchemendy and focus his criticism on our premise (2).

The criticism in a nutshell is as follows. Premise 2 fails because it fails to account for the truth-preservation guarantee, i.e., (2) fails because (i) is false. That there is no possible use of the non-logical terminology of K and X according to which the K-sentences are true and X is false is *a symptom* of the logical guarantee that X is true if the K-sentences are true and not *its cause*. In order for (i) to be true (ii) must be. However, (ii) is false on premise (2)’s reading of *a possible use of language*. To see why (ii) is false consider the following.

(h) (Female(shannon) & ~Married(shannon, matt))

(i) (~Female(matt) & Married(beth, matt))

(j) ~Female(beth)

(j) is not a logical consequence of (h) and (i). By (ii), (j) is not logically guaranteed to be true if both (h) and (i) are true, because the truth-conditional properties of ~ and & permit the existence of a possible use of the predicates and individual constants according to which (h) and (i) are true and (j) is false. However, *the permissibility* of a possible use of the predicates and individual constants according to which (h) and (i) are true and (j) is false is insufficient to assure that there *is, in fact,* such a possible use. To see this consider that if the world contained, say, just two individuals, then there would be no such possible use even though intuitively (j) is still not a logical consequence of (h) and (i), because (j) is not in this situation logically guaranteed to be true if (h) and (i) are true. The reason that the permissibility of a possible use of language does not entail the existence of such a use is the fact that possible uses of language are restricted by what actually exists.

Recall that a possible use of an expression is that expression's potential to possess a semantic value. Further, given a semantic value v, whether an expression has the potential to possess v turns exclusively on the syntactic category assigned to the expression. Obviously, a non-logical expression's potential to possess a certain semantic value depends on the existence of that semantic value (e.g., "Matt" can't refer to Evan if Evan doesn't exist, and "Male" can't be assigned an infinite extension if there aren't infinitely many individuals.) Therefore, that a non-logical expression has the potential to possess a semantic value is a function of both the expression's syntactic category and the real-world existence of the semantic value. Consequently, what uses of language are possible is restricted by what exists (in the world). The complaint here is that the truth-conditional properties of the occurrent logical expressions do not suffice to establish a possible use of the non-logical expressions according to which (h) and (i) are true and (j) is false. To establish that there is such a possible use we require the assumption that there are, in fact, more than two individuals. This illustrates why (ii) is false. More generally, (ii) is false because what exists in the world is a non-logical matter, and possible uses for language are constrained by what exists in the world. Therefore, (i) fails to characterize the source of the truth-preservation guarantee provided by logical consequence. The diagnosis of the problem: Premise (2), from which we derived (i), is false. It is not necessarily the case that it is logically possible for all the K-sentences to be true with X false only if there exists a possible use of the non-logical components of the sentences in set K and of X according to which the former are all true and the later false. In the above scenario, there is no possible use of the non-logical expressions in (h), (i), and (j) according to which (h) and (i) are true and (j) is false, even though intuitively it is logically possible for (h) and (j) to be true and (j) false.

To see the import of the criticism in a different way, consider the sentence '$\forall$x Male(x) $\vee$ $\forall$x~Male(x)'. Clearly, this is not a logical truth. The truth-conditional properties of the logical expressions permit a possible use of 'Male' according to which it is true of one individual and not another. Hence, the sentence is not logically guaranteed to be true. However, the permissibility of this use of the predicate is not sufficient for there being such a possible use of 'Male'. If Parmenides is correct and there is only "The One," then there is no possible use of 'Male' according to which it is true of one individual and not another. In this Parmenidean scenario, '∀x Male(x) ∨ ∀x~Male(x)' would be logically guaranteed to be true. The determination that the denial of '∀x Male(x) ∨ ∀x~Male(x)' is logically possible requires a commitment to the existence of more than one individual. Hence, the fact that this sentence could logically be false not only turns on the semantic functioning of ∀, ∨ , and ~, but also on the fact that there exists at least two individuals.

Of course, there are more than two individuals. The point of the above examples is to explain why (ii) is false by illustrating how what is a logical consequence of what and what is logically true in a first-order language varies across different assumptions regarding the cardinality of the world's individuals, even though the truth-conditional properties of the logical constants is fixed relative to such assumptions. Therefore, the truth-conditional properties of the logical constants do not suffice to provide the assurance that there is no possible use of the non-logical components of the K-sentences and sentence X according to which the former are true and the latter is false. Intuitively, how many individuals there are is a fact that has nothing to do with logic, because it is a fact that has nothing to do with the semantics (i.e., the truth-conditional properties) of the logical expressions of the language. We can't derive how many individuals there are just from an account of the truth-conditional properties of the connectives, identity and the quantifiers.

Let's summarize the challenge raised here to premise (2) in the argument for the adequacy of the model-theoretic characterization of logical consequence. It can't be the case that even though a sentence X isn't a logical consequence of a set K of sentences, X could be. The truth-preservation guarantee requires that on the supposition that the world contains less the extension of logical consequence should not expand. This raises a problem with premise (2), because the range of possible uses for a language varies under different assumptions as to the cardinality of the world's individuals, and the extension of logical consequence inherits this variability. For example, (j) is not, in fact, a logical consequence of (h) and (i), but if (2) is correct, it would be if there were just two things. However, on the suppositition that there are just two individuals, the truth of (h) and (i) do not logically guarantee the truth of (j). The fact that (2) doesn't reflect the truth-preservation guarantee (i.e. because (i) is false) is evidence that it has failed to capture the modal notion inherent in the common concept of logical consequence.

The criticism to premise (2) shows that (ii) is out of sync with our account of a possible use of language. Proposition (ii) needs revision in order to reflect that possible uses of language are constrained by the way the world actually is. We revise (ii) in terms of (iiR).

(iiR) the truth-conditional properties of the occurrent logical expressions *plus* the way the world actually is provide the assurance that there is no possible use of the non-logical components of the sentences in set K and in sentence X according to which the former are true and the latter is false.

(iiR) encapsulates what is wrong with premise 2 according to above criticism: there being no possible use of the non-logical components of the sentences in a set K and in X according to which the former are true and the latter is false cannot be the source of the truth-preservation guarantee, because whether or not there is such a use turns in part on the way the world actually is (e.g., turns, in part, on the cardinality of the world's individuals). But such non-logical matters of fact cannot be a source of the truth-preservation guarantee.

A possible use of language is a use of its non-logical components according to which every sentence receives a truth-value relative to a state of affairs. The criticism of (2) motivates reconceiving possible uses of language so that they are not restricted by what is actually the case. One way of developing this is to construe the state of affairs appealed to in the characterization of a possible use of language as a logically possible state of affairs. If a use of a language L is permissible by the truth-conditional properties of its logical expressions, then it is a possible use of its non-logical components according to which every sentence receives a truth-value relative to a *logically possible* state of affairs. A logically possible state of affairs is one way the world could logically be. Such a state of affairs is logically possible, because there is a use of language permitted by the truth-conditional properties of the logical expressions according to which sentences are true relative to that state of affairs. In other words: if the supposition that there exists a use of certain sentences according to which they are true relative to a state of affairs does not violate the truth-conditional properties of occurrent logical expressions, then the state of affairs obtains *qua* logically possible state of affairs. Here, it is logically possible for a set K of sentences to be true iff there is a possible use of the non-logical components of the K-sentences according to which they are true in some logically possible world. On this approach, the defense of extensional adequacy of the model-theoretic definition of logical consequence replaces (2) with (2').

(2'):there are is no possible use of the non-logical components of the relevant language which makes all the K-sentences and the negation of X true in some logically possible world iff it is not logically possible for all the sentences in K to be true and X false, this is due to the forms of the sentences, and this is knowable *a priori*.

A Parmenidean world is logically possible, because there is a possible use for language M according to which ∀x∀y(x=y) is true. The supposition that ∀x∀y(x=y) is true does not contradict the truth-conditional properties of the sentence's logical components, and these properties determine that the state of affairs which makes the sentence true contain no more than one individual. Hence, the state of affairs according to which ∀x∀y (x=y) is true is a logically possible one.

Consider the following three sentences.

(k) ∀x~ OlderThan(x,x)
(l) ∀x∀y∀z ((OlderThan(x,y) & OlderThan(y,z))→OlderThan(x,z))
(m) ∃x∀y~OlderThan(y,x)

Sentences (k) and (l) tell us that the *older than* relation is irreflexive and transitive, and (m) says that there is an individual which is a minimal element of *older than*. A minimal element is an individual that is at least as old as everyone else (a relation may have more than one minimal element). Sentence (m) is a logical consequence of (l) and (k) iff there is no possible use of the predicate *OlderThan* according to which it refers to a relation *R* that is irreflexive, transitive, and does not have a minimal element (e.g., in order for *old than* not to have a minimal element it would have to be the case that for each individual x there is an individual y such that y is older than x). Since the extension of an irreflexive, transitive relation that does not have a minimal element must be countably infinite, by premise (2) whether there is a possible use of *OlderThan* according to which (k) and (l) are true and (m) is false depends on the existence of a countably infinite collection of individuals. According to (2), if there is not such an existent infinity, then there is no such possible use and (m) is a logical consequence of (k) and (l).

This reflects the idea that whether a use for a language is possible turns on how the world actually is, not how it could have been. By (2'), since a possible use of *OlderThan* according to which its extension is countably infinite is permissible by the truth-conditional properties of the logical components of (k)-(m), an existent infinity is logically possible and so there is a use of *OlderThan* according to which (k) and (l) are true and (m) false in a logically possible world with a countably infinite number of individuals. On this approach, we do not appeal to the way the world is in imposing an upper bound on the cardinality of the totalities of individuals we may consider in determining the possible uses of language. So, (2'), unlike (2), makes possible the invariance of the extension of the logical consequence relation under different considerations as to ways the world could be. Hence, it is coherent to both think that (m) is not a logical consequence of (k) and (l), and believe that the world is finite. The existence of an infinite totality is possible, and relative to such a possible world "OlderThan" may be assigned an infinite extension.

Proposition (2') motivates taking a structure for language M to depict a possible use of it according to which its sentences receive truth-values in a logically possible world, where a logically possible world is one way that the world could logically be. Replacing premise (2) with (2') requires replacing (3) with (3').

(3') there are is no possible use of the non-logical components of the relevant language which makes all the K-sentences and the negation of X true in some logically possible world iff K $\models$X.

A structure with, say, a domain of one individual represents the world with just one individual. Since different domains represent different cardinalities of the world's individuals, we no longer need to take the range of a quantifier to be less than the totality of the world's individuals. According to Etchemendy (p. 287),

> The set-theoretic structures that we construct in giving a model-theoretic semantics are meant to be mathematical models of logically possible ways the world, or relevant portions of the world, might be or might have been. They are used to characterize how variations in the world affect the truth-values of sentences in the language under investigation. ((2008), p. 287)

The class of structures for a language represents all logically possible ways the world might be that are relevant to the truth or falsity of sentences in the language. A sentence S should be true in a model A iff S would be true if the world were depicted by A, that is if A were an accurate model. Any individual model represents a logically possible configuration of the world. (See Shapiro (2005a), p. 663ff, for statements of this view.)

Proposition (2') is allegedly an advance over (2) in grounding the truth-preservation guarantee, because by appealing to logically possible worlds in accounting for logical possibility, the truth-preservation guarantee does not turn on non-logical matters of fact. In order to elaborate, consider the following two propositions.

(i') if there is no possible use of the non-logical components of the sentences in K and of X, according to which the K-sentences are true and X is false in some logically possible word, then X is logically guaranteed to be true if all the K-sentences are true.

(ii') the truth-conditional properties of the occurrent logical expressions provide the assurance that there is no possible use of the non-logical components of K and X according to which the former are true and the latter is false in a logically possible world.

From (2'), we derive (i'), which is defended by (ii').

In response, the notion of a logically possible world that is at work in (i') and (ii') is somewhat problematic. We take logical possibility to apply to sentences and not to non-linguistic states of affairs. It is logically possible that the sentence "Barack Obama is a giraffe" be true, but it is hard to know what it make of the claim that Barak Obama being a giraffe is logically possible even though it may be metaphysically impossible. More importantly, that a use of non-logical expressions is permissible by the truth-conditional properties of logical expressions does not suffice to establish the existence, modal or otherwise, of semantic values for the non-logical expressions. As mentioned earlier, the possible uses of non-logical expressions are assignment of objects and sets of objects to the language's singular terms and predicates. That these meaning assignments are successful depends on the existence—in some sense—of these objects and sets of objects. Facts about the possible semantic values of expressions are not independent of ontological commitments. The very notion of a possible use or meaning of a term relies on substantive, metaphysical claims, whether these are modal or otherwise.

More precisely, letting α range over individuals or a set of individuals, there is a possible use for a expression t according to which it designates α only if (a) and (b) obtain.

(a) nothing about the semantic functioning of t rules out it referring to α, and
(b) α could exist (if α exists, then (ii) is satisfied).

Only if both (a) and (b) are satisfied can t be used to pick out α. (a) is a semantic claim that is independent of metaphysical commitments and is non-substantive. As an aside, there is a problem with fleshing out (a) as a subjunctive conditional—if α could exist, then α could be the semantic value of t. According to some popular theories, the truth-values of such subjunctive conditionals are dependent on the structure of metaphysical possibility. But the account here needs to understand them as making claims that are independent of the metaphysical modal facts. We leave this underdeveloped here. Back to the main point: there is nothing about the semantic functioning of, say, "Shannon McKeon" qua individual constant or the truth-conditional properties of logical expressions that rules out it referring to an impossible thing like a circular square. But "Shannon McKeon" cannot be used to refer to a circular square or to the greatest prime number because such things cannot exist. So, the claim that a term t can be used to pick out α is partly semantic (a) and partly metaphysical (b). Since, possible uses for linguistic items are limited by possibilities for real change, the possibility of using a predicate to pick out an existent infinity is parasitic on the possibility of an existent infinity.

Accordingly, it is inconsistent for one to maintain that there is necessarily a finite number of objects in the world and hold that (m) is not a logical consequence of (k) and (l), because "OlderThan" could be used to pick out a relation with an infinite extension. Again, whether, "OlderThan" could have had an infinite extension is not independent of whether an infinite number of objects could exist. So, if one holds that such a totality is impossible, one cannot countenance a possible use for "OlderThan" which designates such a totality, even though such a use is not ruled out semantically.

What we are questioning is the idea that if the supposition that there exists a use of certain sentences according to which they are true relative to a state of affairs does not violate the truth-conditional properties of occurrent logical expressions, then the state of affairs obtains. Simply calling the state of affairs logically possible does not allay the concern with the alleged derivation of modal features of the world from possible uses of the non-logical components of language. Against this, what seems to be the case is that what uses of language are possible is determined by features of the world, modal or otherwise. It is preferable not to constrain possible uses of language in terms of the mysterious notion "ways the world could be". This results in a stronger epistemological foundation for our judgments of what logically follows from what since we have more to say about what possible uses there are for non-logical components of language than about the modal features of the world.

In reply, perhaps the proponent of (2') can argue that whether a state of affairs is logically possible is independent of possible uses of language. For example, that the world could logically consist of just one individual is independent of the fact that the variables of a first-order language could be used to range over just one individual. This raises the question of how we know that such a state of affairs is logically possible. Etchemendy says that set-theoretic structures that we construct in giving a model-theoretic semantics are meant to be mathematical models of logically possible ways the world, or relevant portions of the world, might be or might have been. So a model whose domain consists of just one individual represents the possibility of the world consisting of a lone individual. How do we know that could there be just one individual? How do we know that it is possible for there to be just finitely many individuals? Indeed, if we countenance an infinite totality of necessary existents such as abstract objects (e.g., pure sets), then it isn't easy to understand how it is possible for there to be just *n* individuals for finite *n* (for discussion see McGee (1999)). One could reply that while it may be metaphysically impossible for there to be merely finitely many individuals it is nevertheless logically possible and this is relevant to the modal notion in the concept of logical consequence. This reply seems to require the existence of primitive, basic intuitions regarding the logical possibility of there being just finitely many things. However, intuitions about possible cardinalities of worldly individuals—not informed by mathematics and science—tend to run stale. Consequently, it is hard to debate this reply: one either has the needed logical intuitions, or not.

Proposition (iiR), which is the rationale for thinking that premise (2) accounts for the source of the truth-preservation guarantee, reflects the idea that knowledge of possible uses of language is grounded on knowledge of substantive facts about the world. This, in turn, is reflected in our knowledge of the extension of the model-theoretic consequence relation. Our knowledge of what is a model-theoretic consequence of what in a language L depends on our knowledge of the class of L-structures. Since such structures are furniture of the world, our knowledge of the model-theoretic consequence relation is grounded on knowledge of substantive facts about the world. Even if such knowledge is *a priori*, it is far from obvious that our *a priori* knowledge of the logical consequence relation is so substantive. One might argue that knowledge of what follows from what shouldn't turn on worldly matters of fact, even if they are necessary and *a priori*. If correct, this is a strike against the model-theoretic definition. However, this standard logical positivist line has been recently challenged by those who see logic penetrated and permeated by metaphysics (e.g., Putnam (1971), Almog (1989), Sher (1991), Williamson (1999)). We illustrate the insight behind the challenge with a simple example. Consider the following two sentences.

(n)$\exists$x (Female(x) & Sister(x, evan))
(o) $\exists$xFemale(x)

(o) is a logical consequence of (n), i.e., there is no domain for the quantifiers and no interpretation of the predicates and the individual constant in that domain which makes (n) true and not (o). Why? Because on any interpretation of the non-logical terminology, (n) is true just in case the intersection of the set of objects that satisfy *Female(x)* and the set of objects that satisfy *Sister(x, evan)* is non-empty. If this obtains, then the set of objects that satisfy *Female(x)* is non-empty and this makes (o) true. The basic metaphysical truth underlying the reasoning here is that for any two sets, if their intersection is non-empty, then neither set is the empty set. This necessary and *a priori* truth about the world, in particular about its set-theoretic part, is an essential reason why (n) follows from (o). This approach, reflected in the model-theoretic consequence relation (see Sher (1996)), can lead to an intriguing view of the formality of logical consequence reminiscent of the pre-Wittgensteinian views of Russell and Frege. Following the above, the consequence relation from (n) to (o) is formal because the metaphysical truth on which it turns describes a formal (structural) feature of the world. In other words: it is not possible for (n) to be true and (o) false, because for any extensions of P, P', it is the case that if for some object x, it is true that (P(x) & P'(x, evan)), then for some object x, it is true that P(x).

According to this vision of the formality of logical consequence, the consequence relation between (n) and (o) is formal because what is in bold expresses a formal feature of reality. Russell writes that, "Logic, I should maintain, must no more admit a unicorn than zoology can; for logic is concerned with the real world just as truly as zoology. Though with its more abstract and general features" (Russell, (1919) p. 169). If we take the abstract and general features of the world to be its formal features, then Russell's remark captures the view of logic that emerges from anchoring the necessity, formality and a priority of logical consequence in the formal features of the world. The question arises as to what counts as a formal feature of the world. If we say that all set-theoretic truths depict formal features of the world, including claims about how many sets there are, then this would seem to justify making , $\exists x \exists y \sim (x=y)$, (i.e. *there are at least two individuals*) a logical truth since it is necessary, *a priori*, and a formal truth. To reflect model-theoretically that such sentences, which consist just of logical terminology, are logical truths we would require that the domain of a structure simply be the collection of the world's individuals. See Sher (1991) for an elaboration and defense of this view of the formality of logical truth and consequence. See Shapiro (1993) for further discussion and criticism of the project of grounding our logical knowledge on primitive intuitions of logical possibility instead of on our knowledge of metaphysical truths.

To summarize, our case for the extensional adequacy of the model-theoretic characterization of logical consequence calls up on (2).

(2) It is not logically possible for all the sentences in K to be true and X false, this is due to the forms of the sentences, and this is knowable *a priori* iff there is no possible use of the non-logical components of the language according to which all of the sentences in K are true and X is false.

Using (2), we explain the source of the truth-preservation guarantee in terms of (i), which is supported by (iiR).

(i) if there is no possible use of the non-logical components of the sentences in K and of X, according to which the K-sentences are true and X is false, then X is logically guaranteed to be true if all the K-sentences are true.

(iiR) the truth-conditional properties of the occurrent logical expressions *plus* the way the world actually is provide the assurance that there is no possible use of the non-logical components of the sentences in set K and in sentence X according to which the former are true and the latter is false.

The criticism of (2) is essentially that (iiR) is inadequate because it makes the way the world actually is, a substantive, metaphysical matter, a source of the truth-preservation guarantee. If cogent, the criticism motivates replacing (2) with (2').

(2') There are is no possible use of the non-logical components of the relevant language which makes all the K-sentences and the negation of X true in some logically possible world iff it is not logically possible for all the sentences in K to be true and X false, this is due to the forms of the sentences, and this is knowable *a priori*.

From (2') we derived (i'), which is an alternative to (i) as an explanation of the source of the truth-preservation guarantee. (ii') is offered as a reason for (i').

(i') if there is no possible use of the non-logical components of the sentences in K and of X, according to which the K-sentences are true and X is false in some logically possible word, then X is logically guaranteed to be true if all the K-sentences are true.

(ii') the truth-conditional properties of the occurrent logical expressions provide the assurance that there is no possible use of the non-logical components of K and X according to which the former are true and the latter is false in a logically possible world.

For those of us who are skeptical that there are logically possible worlds (i.e., ways the world could logically be) there is reason to reject (ii') and doubt that (i') makes the source of the truth-preservation guarantee any less of a substantive, metaphysical matter than (i). Whether the sentences of a language L can be used to depict a state of affairs, modal or otherwise, depends on whether the state of affairs obtains, and this is independent of any facts regarding the semantic functioning of the expressions of L. Studying the truth-conditional properties of language M's quantifiers as reflected in the quantifier truth clauses is not going to tell us whether the world could have evolved so that it consisted of just one individual. To say that such a state of affairs is logically possible either ignores what we take to be the correct account of possible uses for language or mystifies how we know that a state of affairs is logically possible even if metaphysically impossible. If the appeals to logically possible worlds in (i') and (ii') are replaced with appeals to metaphysically possible worlds, then (ii') is obviously false and (i') thus construed undercuts the criticism that (i) is defective because it makes metaphysical matters a source of the truth-preservation guarantee.

Since there are obvious epistemological advantages to making whether a given use of language is possible turn on whether actual rather than merely modal states of affairs obtain, it has to be clear why the appeal to possible worlds in (i') makes it preferable to (i) in explaining the source of the truth-preservation guarantee. Since this hasn't been made clear, there is reason to think that (iiR) is defensible and that the criticism of premise 2 is not successful.

Premise (3): Are there possible uses not captured by possible interpretations?

The third premise in support of the extensional adequacy of the model-theoretic characterization of logical consequence is as follows.

(3) There is no possible use of the non-logical components of the language according to which all of the sentences in set K are true and X is false iff K $\models$X.

Premise (3) says in effect that the model-theoretic characterization of logical consequence for M is extensionally equivalent with its characterization in terms of possible uses of M. Premise (3) is an essential link between $\models$ and logical consequence. If true, (3) makes the fact that the K-sentences and negated X are true relative to a structure evidence that X is not a logical consequence of K. In other words: by the lights of premise (3), there being no structure according to which the K-sentences are true and X is false suffices for X being a logical consequence of K, because the class of structures represents all the possible uses for the relevant language.

Clearly, if there is a model of a set of sentences (i.e., an interpretation that makes all the sentences true relative to a domain), then there is a possible use of the non-logical elements of the relevant language according to which the sentences are true. This makes the if-part of (3) plausible (i.e., makes plausible that if there is no possible use of the non-logical components of the language according to which all of the sentences in set K are true and X is false, then K $\models$X). Towards the evaluation of the only-if-part of (3), is it true that if there is a possible use of the non-logical elements of the relevant language according to which each member of a set of sentences is true, then there is a model of the sentences? The above discussion of indefinite extensibility suggests that possible uses outrun models, at least as standardly construed. Intuitively, '∀x(x=x)' can be used to say that absolutely everything is self-identical, '∀x∃y Greater(y,x)' can be used to say that for each ordinal there is a greater one. But standard models cannot represent such uses of the quantifiers, since *all things* and *all ordinals* do not form domains. Of course, this doesn't suggest that there is a model of the negation of the first sentence, or a negation of the second in the collection of all ordinals.

Is there a set K of sentences and sentence X from a first-order language L such that K ⊨X even though there is a possible use of L according to which the K-sentences are true and X is not true? To sharpen the question, premise (3) requires the truth of (3').

(3'): for every possible use of the non-logical components of the relevant language which makes all the K-sentences and the negation of X true, there is a model of K and the negation of X.

Is (3') true? There is reason to think that it is false, and this motivates criticism of (3). We start our discussion of (3') by noting that our model-theoretic characterization of logical consequence reflects what is referred to as classical model-theoretic semantics, since for every M-structure U (i.e., every structure that interprets our language M) the following obtain.

(i) A constant of M refers to an element of U's domain.

(ii) The domain D is non-empty, and so M's variables always range over a non-empty domain.

(iii) An n-place predicate designates a set of n-tuples that is a subset of D^n.

To illustrate (iii), a one-place predicate P designates a subset S of D, i.e., the extension of P restricted to D. We spell out the semantic details with respect to one-place predicates. Each element of S is what P is true of; each element of $\overline{S}$, the complement of S restricted to D, is what P is false of. Hence, under an interpretation of P relative to a domain D, P divides D into the subset S of elements P is true of and the subset $\overline{S}$ of elements D is false of ($S \cup \overline{S} = D$). Also, no predicate P is both true of and false of an object. That is, the intersection of the subset of domain D that is P's extension and the subset of elements that P is false of is empty ($S \cap \overline{S} = \varnothing$).

In what follows, we generate criticism of (3') by questioning the rationale of (i)-(iii). Of course, questioning (i)-(iii) does not question the use of structures to fix what follows logically from what. Rather, the issues are whether classical model-theoretic semantics is correct and whether alternatives are motivated. In the space allotted for consideration of premise (3) we cannot do justice to the substantive issues raised in considering the rationale for (i)-(iii). Our primary aim here is to develop enough criticism of premise (3) to illustrate both that the truth-conditional properties of the logical expressions are (partly) derived from the semantics of terms and predicates, and relatedly that the model-theoretic definition of logical consequence is not an analysis of logical consequence.

Consider the following three sentences from our language M.

(p) Male(evan)
(q) $\forall$x male(x)
(r) $\exists$x Male(x)

Sentence (r) is a model-theoretic consequence of (p) because (i) obtains and it's a model-theoretic consequence of (q) because (ii) obtains. What is the rationale for (i) and (ii)? Earlier, we said that variations in the semantic values of individual constants, predicates, and variables represent possible uses of the relevant language. Recall that we regard an alternative possible use/meaning of an expression as the expression's potential—relative to its assignment in a syntactic category—to have that use/possess that meaning.

In this context, in order to motivate (i) and (ii) we must say (as was said earlier in this chapter) that there is no possible use of variables that makes them range over an empty domain and no use of an individual constant according to which it lacks a referent. But these claims may be reasonably questioned. Why can't a language be used so that some of its individual constants lack a referent or used so that its variables range over the empty domain? Even if it is true that the ordinary-language correlates of M's individual constants and variables are not in fact used this way, this is insufficient to rule out the possibility of such uses. Concerning language M, it certainly seems that an individual constant has the semantic potential to lack a referent and that variables have the potential to range over an empty domain. It is hard so see how this could be questioned. Certainly, the metaphysical essences of a constant and variable do not rule out such uses. What uses are possible for these linguistic entities is a matter of convention.

By virtue of depicting possible uses for language, structures are representations of logical possibilities. By considering the widest range of possibilities in fixing what follows logically from what, we underwrite the truth-preservation guarantee of the logical consequence relation. Plausibly, this motivates taking possible uses for individual constants and variables to range as widely as possible. By stipulating that variables have the semantic potential to range over the empty set and individual constants have the semantic potential to lack a referent, we make such uses possible, and, therefore, they are relevant to determining the logical consequence relation. Therefore, (r) should not be a model-theoretic consequence of (q), because if (q) is true, then (r) is not guaranteed to be true. The variables in (q) and (r) can be used to range over the empty domain, and when used this way (q) is true and (r) is false. Clearly, the legitimacy of this use is independent of the issue of whether the universe could (in a metaphysical sense) have been empty. Also, given that 'evan' can be used so that is fails to refer, (r) should not a model-theoretic consequence of (p) unless we add '∃x (x=evan)' to (p).

If we grant uses of constants according to which they fail to refer, then how do we evaluate atomic sentences containing constants used in such a way? Consider the following three sentences.

(s) Frodo is a hobbit
(t) Frodo is an elf
(u) Frodo has a tattoo on his belly

Intuitively, (s) is true, (t) is false, and (u) is neither (since it is not determined by the Tolkien stories). Admitting atomic propositions that are neither true nor false obviously impacts what the truth-conditional properties of the logical expressions are. Indeed, the semantics of atomic wff constrains the correct account of the truth-conditional properties of logical expressions.

For an example of how the semantics of atomic sentences drives the account of the truth-conditional properties of logical expressions, suppose that 'A' is an atomic sentence that is neither true nor false, and 'B' is a true atomic sentence. What is the truth-value of 'A v B'? Does 'A's lack of a truth-value transfer to 'A v B' or is the disjunction true because 'B' is? This is a question about the truth-conditional properties of 'v', and, therefore, a question about the meaning of 'v' in this context. For illustrative purposes, let's say that a disjunction is false iff either both disjuncts are false or one disjunct is false and the other lacks a truth-value. A disjunction is true iff when at least one disjunct is true. According to classical semantics, 'A v ~A' is true for every sentence A, and so is a pattern of a logical truth. But this isn't the case on the proposed understanding of 'v' since it isn't true when 'A' lacks a truth-value. In such a case, since 'A' lacks a truth-value, '~A' does as well. One may appeal to the notion of a supervaluation and grant the possibility of uses of individual constants according to which they fail to refer while preserving the classical semantic evaluation of those sentences that are logically true solely due to the truth-conditional properties of the sentential connectives. We follow van Fraassen (1969).

Let L be a language without identity that contains just atomic sentences and the sentential connectives. A partial evaluation of L is one that does not assign a truth-value to every atomic sentence. A classical extension of a partial evaluation is an extension of the partial assignment that arbitrary assigns truth-values to those sentences that received none according to the partial assignment. Relative to a partial assignment for L, a *supervaluation* for L assigns "true" to those L-sentences that are true on every classical extension, "false" to those sentences that are false on every classical extension, and doesn't assign a truth-value to the sentences that are true on some extensions and false on others. A supervaluation then either assigns a truth-value to a compound sentence containing atomic components which lack a truth-value that value that all classical valuations would assign it if there is a such value, or the supervaluation assigns it no value. A sentence is a logical truth if it is true on every supervaluation, i.e., there is no classical extension of a partial evaluation according to which it is false. A sentence X is a logical consequence of a set K of sentences iff there is no supervaluation according to which the K-sentences are true and X is false (i.e., there is no partial evaluation such that every classical extension of it makes the K-sentences are true and X false).

Consider again, 'A v ~A', where 'A' lacks a truth-value according to some partial evaluation. One extension makes 'A' true, and another makes it false. Employing the classical accounts of the truth-conditional properties of 'v' and '~', the disjunction is true on both extensions. Since this is the only partial evaluation relevant to the evaluation of 'A v ~A' it follows that the disjunction is true on every supervaluation, and is a logical truth as defined above. Note that as a consequence, 'v' is no longer truth-functional since 'A v ~A' is true when both disjuncts lack a truth-value ('~A' lacks a truth value if 'A' does), but 'A v B' lacks a truth-value when both disjuncts do (e.g., 'Frodo has a tattoo on his belly v ~Frodo has a tattoo on his belly' is a logical truth, but 'Frodo has a tattoo on his belly v Frodo has a cavity in his right bicuspid' lacks a truth-value even though just like the first disjunction both disjuncts lack a truth-value).

The logic that arises from dropping (i) and (ii) is called "free logic", a logic free from existential assumptions. There is certainly much more to say regarding free logic. In Chapter 5, we discuss free logic in connection with the deductive system developed there. What has been illustrated here is that whether or not (r) is a logical consequence of (p) and of (q) turns on the semantic constraints (i) and (ii), which concern the semantics of variables and constants. More broadly, the very truth-conditional characteristics of logical expressions such as the sentential connectives turns on the semantics for the non-logical terminology. This suggests that it is circular to argue that, say, since (r) is a logical consequence of (p), (i) obtains. Rather, whether or not (i) obtains is an issue prior to the issue of whether (r) is a logical consequence of (p) and must be decided beforehand. Similarly, what the correct semantics is for variables and constants constrains adequate accounts of the truth-conditional properties of logical expressions. A similar theme is played out in our discussion of (iii), which we now begin.

For any M-sentence α, both the following hold.

$\models (\alpha \vee \sim \alpha)[g_{\varnothing}]$.
$\models \sim (\alpha \,\&\! \sim \alpha)[g_{\varnothing}]$

The first, mentioned above, reflects what is sometimes called the law of excluded middle, and the second the law of contradiction. Given that the satisfaction clauses in the truth definition for M are recursive (i.e., the satisfaction of a compound wff is derived from the satisfaction of component atomic wffs), the two laws require that **U** $\models R(t_1, \ldots, t_n)[g]$ or **U** $\not\models R(t_1, \ldots, t_n)[g]$ (i.e., by II, **U** $\models \sim R(t_1, \ldots, t_n)[g]$). Recall the first satisfaction clause in the truth definition for M: **U** $\not\models R(t_1, \ldots, t_n)[g]$ iff the *n*-tuple $< t_1[g], \ldots, t_n[g]> \notin I_U(R)$. Hence, the laws of excluded middle and contradiction, as understood above, require that either $< t_1[g], \ldots, t_n[g]> \in I_U(R)$ or $<t_1[g], \ldots, t_n[g]> \notin I_U(R)$. This necessitates that for any property P defined on a domain D (v) it is determinate what P is true of, and also (w) nothing from D is both true of P and not true of P. If a property P fails (v), we say that it is vague. If it fails (w), we say it is contradictory. So, (A) requires that a property P defined on a domain be neither vague nor contradictory. This is basically the requirement, discussed above, that P supply a definite criterion of application.

What is the motivation for (v) and (w)? That vague or contradictory properties are not possible is a metaphysical thesis that has been seriously questioned. Since vagueness was previously discussed in this Chapter, here we focus on (w). Recent developments in paraconsistent logic make it unlikely that there is a formal reason for maintaining (w). A paraconsistent logic is defined as a logic according to which explosion fails (i.e., for all sentences P, Q, Q is not a logical consequence of P & ~P). Borrowing from Beall (2003), we introduce the propositional fragment of a paraconsistent logic.

Let L be a language without identity that contains just atomic sentences and the sentential connectives. Interpretations of L are functions from L-sentences into the powerset of {T,F}. The powerset of a set S is the set of S's subsets. The powerset of {T,F} is {{T,F}, {T}, {F}, ∅}. Hence, where P is any sentence, V(P) may be any member of the powerset. Intuitively, the following obtain.

'T ∈ V(P)' means that P is at least true according to interpretation V.
'F ∈ V(P)' means that P is at least false according to V.
'∅ =V(P)' means that P is neither true nor false according to V.
'T ∈ V(P) and F ∈ V(P)' means that P is true and false according to V.

The truth-conditional properties of the connectives are determined as follows.

(a) T ∈ V(¬P) iff F ∈ V(P)
(b) F ∈ V(¬P) iff T ∈ V(P)

(c) T∈ V(P & Q) iff T ∈ V(P) and T ∈ V(Q)
(d) F ∈ V(P & Q) iff F ∈ V(P) or F∈ V(Q)

(e) T∈ V(P v Q) iff T ∈ V(P) or T ∈ V(Q)
(f) F∈ V(P v Q) iff F∈ V(P) and F∈ V(Q)

A sentence X is a logical consequence of a set K of sentences iff there is no interpretation according to which all the sentences in K are *at least* true and X is false (i.e., for every interpretation V, if the K-sentences are at least true with respect to V, then X is at least true according to V). For example, P v Q is a logical consequence of {P}.

Proof: Suppose that P is at least true on some interpretation , i.e., suppose that T ∈ V(P), for some interpretation V. Then by (e), T∈ V(P v Q).

Q is not a logical consequence of {P, ¬P v Q}.

Proof: Consider the following interpretation. V(P)={T,F}, V(Q)={F}. Then (i) T ∈ V(P) and (ii) F ∈ V(P). From (ii) and (a), (iii) T ∈ V(¬P). From (iii) and (e), (iv) T∈ V(¬P v Q). This shows that the two premises are at least true according to V (from (i) and (iv)), and the conclusion is false according to V (V(Q)={F}).

Q is not a logical consequence of {P ∧ ¬P} (Explosion fails).

Proof: Consider the following interpretation. V(P)={T,F}, V(Q)={F}. Here Q is false and P is at least true and at least false. Fairly straightforward to work out (by appealing to (b) above) that the contradiction is a least true and the conclusion false according to V.

A *dialethia* is any true statement of the form *P and it is false that P*. *Strong Paraconsistentism* is the view that there are logical possibilities according to which contradictions are true; however, no contradiction is in fact true. *Dialethic Paraconsistentism* holds that some contradictions are in fact true, and so entails strong paranconsistentism. In order to sketch a rationale for rejecting the law of non-contradiction, we consider one offered by Graham Priest, a well-known dialethic paraconsistentist.

Consider the following two conception of infinitude. An *actually infinite* collection is an infinite collection all of whose members are given at one point in time (the infinitude of the collection is given all at once). A *potentially infinite* collection is a finite collection that always can be extended, i.e., that always can be added to with unique objects of the same kind. The typical advocate of the existence of potentially infinite collections (e.g., Aristotle) maintains that infinitude is never wholly present, and thus rejects that there are actually infinite collations.

The extensions of indefinitely extensible properties are potentially infinite collections. We have said that, for example, no collection corresponds to the unbounded increasing sequence of ordinals. However, according to the domain principle, for every potential infinity there exists a collection of exactly the members of the potential infinity. In essence, the domain principle rejects the existence of non-collectivizing properties.

Priest accepts the domain principle. He thinks that it is a brute semantic fact that whenever there are things of a certain kind, there are *all* of those things. ((2002), p.280) In particular, Priest thinks that for any claim of the form ‘all Ss are Ps’ to have determinate sense there must be a determinate totality over which the quantifier ranges. (Ibid, p. 125ff) For example, in the above characterization of the ordinal progression, and in our subsequent discussion of it there are numerous ostensibly universal quantifications of roughly the form 'All ordinals are such and such'. The following are just some of the many instances.

(i) For all ordinals α, there exists an ordinal $\alpha +1$ which is the immediate successor of α.
(ii) Let each ordinal be identified with the collection of preceding ordinals.
(iii) The ordinal progression is not the collection of all ordinals.

Intuitively, these sentences are about *all* ordinals, i.e., they speak of a collection of ordinals. If (i) and (ii) are meaningful, then there is a determinate totality (of ordinals) each of whose members they characterize in some way. Furthermore, it seems that in apprehending the meaning of (iii), we conceptualize both the ordinal progression and the collection of all ordinals as unities, i.e., as collections given all at once. Following Priest, in asserting (iii) we are saying that the collection of all ordinals, *that totality*, is not identical with the totality that corresponds with the ordinal progression. On this reading, Priest regards (iii) as false.

In essence, the domain principle rejects the existence of non-collectivizing properties. There is a collection of all ordinals, i.e., what *being an ordinal* is true of forms a collection, which Priest takes the ordinal progression to characterize. We now work off of our previous discussion of ordinals and the concept of indefinite extensibility.

The collection of all ordinals, ON, is an ordinal, since ON is the least ordinal greater than every element in the collection of all ordinals. But then ON is not a member of this collection (since it is not greater than itself), and so ON is not an ordinal. (Ibid, p. 119). From commonly accepted principles regarding ordinals, and Priest's claim that *being an ordinal* is collectivizing, it follows that ON is both an ordinal and not an ordinal. On this view, the property of *being an ordinal* is a contradictory property, i.e., a property that is true of an object and not true of it. Indeed, on this view any indefinitely extensible sortal property is a contradictory property.

If correct, then there are possible uses for predicates according to which they designate contradictory properties. This motivates paraconsistent logic and the abandonment of the law of non-contradiction. Of course, Priest may be wrong. There is criticism of both the domain principle and the inferences Priest draws from it (see Priest's review in (2002) Chapter 17, for discussion and references). One moral to draw from this all-too-short discussion of paraconsistent logic is that an answer to the metaphysical question of whether there are contradictory properties is prior to and thus a determinant of the correct account of the truth-conditional properties of the logical expressions from which the identification of logical laws is derived. It is entirely question-begging to argue against there being contradictory properties on the basis of the truth of the law of non-contradiction. The logic of the debate must not presuppose what is at issue.

We now conclude by drawing a lesson from our discussion of premise (3) regarding the status of the model-theoretic characterization. It should not be regarded as an analysis of the concept of logical consequence, because its use to fix what is logically true and what follows logically from what presupposes such an analysis. For example, if we have serious doubts about whether the law of excluded middle or the law of non-contradiction are true, or doubts about whether '∃xMale(x)' is a logical consequence of 'Male(evan), classical model-theoretic semantics will not diminish them. For the same intuitions that suggest these are logical laws and that inform what follows logically from what are used in defining the class of structures that are employed by this semantics.

We should, therefore, agree with Etchemendy that we cannot look to model-theoretic semantics to answer the most basic foundational issues in logic, since model theory properly understood, does not yield an analysis of logical properties such as logical truth and consequence but presupposes them ((2008), p.265ff). This raises the question: how exactly does the model-theoretic characterization engender understanding of logical concepts such as logical truth and consequence? Well, each structure in the class of structures for a language L depicts a possible use for L. A model-theoretic semantics shows how the truth-values of sentences in L vary as the uses of the language represented by structures vary, and, accordingly, explains the truth-preservation guarantee of the logical consequence relation, and explains the persistence of the truth of logically true sentences due to the semantics of the logical terminology.

This explanation relies on an understanding of what structures depict as representations of logically possible situations as this is reflected in premise (2) in the argument for the extensional adequacy of the model-theoretic characterization. Assuming such an understanding, model-theoretic semantics shows precisely how the logic of the language arises from the truth-conditional properties of the logical expressions. Classical model-theoretic semantics may not explain why the law of non-contradiction is, but it does explain why, given this basic assumption, (P &~ (Q & R)) v Q) is a logical consequence of P. This is enlightening.

What is a Logical Constant?

We now turn to the issue of what qualifies as a logical constant. Note that model theory by itself does not provide the means for drawing a boundary between the logical and the non-logical. Indeed, its use presupposes that a list of logical expressions is in hand. The point here is that the use of models to capture the logical consequence relation requires a prior choice of what terms to treat as logical. This is, in turn, reflected in the identification of the terms whose interpretation is constant from one structure to another. Tarski writes,

> No objective grounds are known to me which permit us to draw a sharp boundary between [logical and non-logical terms]. It seems possible to include among logical terms some which are usually regarded by logicians as extra-logical without running into consequences which stand in sharp contrast to ordinary usage. (1936) p. 418-419

And at the end of his (1936), he tells us that the fluctuation in the common usage of the concept of consequence would be accurately reflected in a relative concept of logical consequence, i.e. a relative concept "which must, on each occasion, be related to a definite, although in greater or less degree arbitrary, division of terms into logical and extra logical" (p.420). Unlike the relativity described in the above discussion of premise (1) in the argument for the adequacy of the model-theoretic characterization which speaks to the features of the concept of logical consequence, the relativity contemplated by Tarski concerns the selection of logical constants.

Tarski's observations of the common concept do not yield a sharp boundary between logical and non-logical terms. It seems that the sentential connectives and the quantifiers of our language M about the McKeons qualify as logical if any terms of M do. We've also followed many logicians and included the identity predicate as logical (See Quine (1986) for considerations against treating '=' as a logical constant). But why not include other predicates such as 'OlderThan'?

(1) OlderThan(kelly,paige)
(2) ~OlderThan(paige,kelly)
(3) ~OlderThan(kelly, kelly)

Then the consequence relation from (1) to (2) is necessary, formal, and *a priori* and the truth of (3) is necessary, formal and also *a priori*. If treating

'OlderThan' as a logical constant does not do violence to our intuitions about the features of the common concept of logical consequence and truth, then it is hard to see why we should forbid such a treatment. By the lights of the relative concept of logical consequence, there is no fact of the matter about whether (2) is a logical consequence of (1) since it is relative to the selection of 'OlderThan' as a logical constant. On the other hand, Tarski hints that even by the lights of the relative concept there is something wrong in thinking that *B* follows from *A and B* only *relative* to taking 'and' as a logical constant. Rather, *B* follows from *A and B* we might say absolutely since 'and' should be on everybody's list of logical constants. But why do 'and' and the other sentential connectives, along with the identity predicate and the quantifiers have more of a claim to logical constancy than, say, 'OlderThan'? Tarski (1936) offers no criteria of logical constancy that help answer this question.

On the contemporary scene, there are three general approaches to the issue of what qualifies as a logical constant. One approach is to argue for an inherent property (or properties) of logical constancy that some expressions have and others lack. For example, topic neutrality is one feature traditionally thought to essentially characterize logical constants. The sentential connectives, the identity predicate, and the quantifiers seem topic neutral; they seem applicable to discourse on any topic. The predicates other than identity such as 'OlderThan' do not appear to be topic neutral, at least as standardly interpreted, e.g., 'OlderThan' has no application in the domain of natural numbers. We noted in Chapter 2 that Tarski in his (1941) logic text identifies "topic-neutrality" as essential characteristic of logical expressions, which reflects their indispensability as a "means for conveying human thoughts and for carrying out inferences in any field whatsoever." The topic-neutrality approach to logical constancy is reflected in many if not most introductory logic texts, and has historical precedence over the two other approaches discussed below.

One way of making the concept of topic neutrality precise is to follow Tarski's suggestion in his (1986) that the logical notions expressed in a language L are those notions that are invariant under all one-one transformations of the domain of discourse onto itself. A one-one transformation of the domain of discourse onto itself is a one-one function whose domain and range coincide with the domain of discourse. And a one-one function is a function that always assigns different values to different objects in its domain (i.e., for all x and y in the domain of f, if $f(x)=f(y)$, then $x=y$).

Consider ‘OlderThan’. By Tarski’s lights, the notion expressed by the predicate is its extension, i.e. the set of ordered pairs <d, d'> such that *d* is older than *d'*. Recall that the extension is:

{<Beth, Matt>, <Beth, Shannon>, <Beth, Kelly>, <Beth, Paige>, <Beth, Evan>, <Matt, Shannon>, <Matt, Kelly>, <Matt, Paige>, <Matt, Evan>, <Shannon, Kelly>, <Shannon, Paige>, <Shannon, Evan>, <Kelly, Paige>, <Kelly, Evan>, <Paige, Evan>}.

If ‘OlderThan’ is a logical constant its extension (the notion it expresses) should be invariant under every one-one transformation of the domain of discourse (i.e. the set of McKeons) onto itself. A set is invariant under a one-one transformation *f* when the set is carried onto itself by the transformation. For example, the extension of ‘Female’ is invariant under *f* when for every *d, d* is a female if and only if *f(d)* is. ‘OlderThan’ is invariant under *f* when <d, d'> is in the extension of ‘OlderThan’ if and only if <f(d), f(d')> is. Clearly, the extensions of *Female* and the *OlderThan* relation are not invariant under every one-one transformation. For example, Beth is older than Matt, but *f*(Beth) is not older than *f*(Matt) when *f*(Beth)=Evan and *f*(Matt)=Paige. Compare the identity relation: it is invariant under every one-one transformation of the domain of McKeons because it holds for each and every McKeon. The invariance condition makes precise the concept of topic neutrality. Any expression whose extension is altered by a one-one transformation must discriminate among elements of the domain, making the relevant notions topic-specific.

The invariance condition can be extended in a straightforward way to the quantifiers and sentential connectives (see McCarthy (1981) and McGee (1997)). Here I illustrate with the existential quantifier. Let Ψ be a well-formed formula with ‘x’ as its free variable. $\exists x\Psi$ has a truth-value in the intended structure U^M for our language M about the McKeons. Let *f* be an arbitrary one-one transformation of the domain D of McKeons onto itself. The function *f* determines an interpretation I' for Ψ in the range D' of *f*. The existential quantifier satisfies the invariance requirement for $U^M \models \exists x\Psi$ if and only if $U \models \exists x\Psi$ for every U derived by a one-one transformation *f* of the domain D of U^M (we say that the U’s are isomorphic with U^M).

For example, consider the following existential quantification.

$$\exists x\ \text{Female}(x)$$

This is true in the intended structure for our language M about the McKeons (i.e., $U^M \models \exists x\ Female(x)[g_0]$) ultimately because the set of elements that satisfy 'Female(x)' on some variable assignment that extends g_0 is non-empty (recall that Beth, Shannon, Kelly, and Paige are females). The cardinality of the set of McKeons that satisfy an M-formula is invariant under every one-one transformation of the domain of McKeons onto itself. Hence, for every U isomorphic with U^M, the set of elements from D^U that satisfy 'Female(x)' on some variable assignment that extends g_0 is non-empty and so the following obtains.

$$\mathbf{U} \models \exists x Female(x)\ [g_0].$$

Speaking to the other part of the invariance requirement given two paragraphs back, clearly for every U isomorphic with U^M, if $U \models \exists x Female(x)\ [g_0]$, then $U^M \models \exists x Female(x)\ [g_0]$ (since U is isomorphic with itself). Crudely, the topic neutrality of the existential quantifier is confirmed by the fact that it is invariant under all one-one transformations of the domain of discourse onto itself.

Key here is that the cardinality of the subset of the domain D that satisfies an L-formula under an interpretation is invariant under every one-one transformation of D onto itself. For example, if at least two elements from D satisfy a formula on an interpretation of it, then at least two elements from D' satisfy the formula under the I' induced by *f*. This makes not only 'All' and 'Some' topic neutral, but also any cardinality quantifier such as 'Most', 'Finitely many', 'Few', 'At least two' etc… The view suggested in Tarski (1986, p. 149) is that the logic of a language L is the science of all notions expressible in L which are invariant under one-one transformations of L's domain of discourse. For further discussion, defense of, and extensions of the Tarskian invariance requirement on logical constancy see McCarthy (1981), McGee (1997), and (Sher (1989), (1991)).

A second approach to what qualifies as a logical constant is to not make topic neutrality a necessary condition for logical constancy. This undercuts at least some of the significance of the invariance requirement. Instead of thinking that there is an inherent property of logical constancy, we allow the choice of logical constants to depend, at least in part, on the needs at hand, as long as the resulting consequence relation reflects the essential features of the intuitive, pre-theoretic concept of logical consequence. I take this view to be very close to the one that we are left with by default in Tarski (1936). The approach, suggested in Prior (1976) and developed in related but different ways in Hanson (1996) and Warmbrod (1999), amounts to regarding logic in a strict sense and loose sense. Logic in the strict sense is the science of what follows from what relative to topic neutral expressions, and logic in the loose sense is the study of what follows from what relative to both topic neutral expressions and those topic centered expressions of interest that yield a consequence relation possessing the salient features of the common concept.

Finally, a third approach to the issue of what makes an expression a logical constant is to simply reject the view of logical consequence as a formal consequence relation, thereby nullifying the need to distinguish logical terminology in the first place (see Etchemendy (1983) and the conclusion in (1999b), and see Bencivenga (1999)). We just say, for example, that X is a logical consequence of a set K of sentences if the supposition that all of the K are true and X false violates the meaning of component terminology. Hence, 'Female(kelly)' is a logical consequence of 'Sister(kelly, paige)' simply because the supposition otherwise violates the meaning of the predicates. Whether or not 'Female' and 'Sister' are logical terms doesn't come into play.

This concludes our discussion of the status of the model-theoretic characterization of logical consequence. We now complete our presentation of a full logic for language M by offering an account of the deductive or proof-theoretic consequence relation in M.

Chapter 5
Deductive Consequence

In defining logical consequence for our language M in terms of the deductive consequence relation, we appeal to a natural deductive system that originates in the work of the mathematician Gerhard Gentzen (1934) and the logician Fredrick Fitch (1952). We will refer to the deductive system as N (for ‘natural deduction’). For an in-depth introductory presentation of a natural deductive system very similar to N see Barwise and Etchemendy (2002). N is a collection of inference rules. A proof of X from K that appeals exclusively to the inference rules of N is a formal deduction or formal proof. A formal proof relative to a language L is associated with a pair <K, X> where K is a set of sentences from L and X is an L-sentence. The set K of sentences is the basis of the deduction, and X is the conclusion. We shall spend some time in what follows later explaining the nature of a formal deduction. But for now, we say that a formal deduction from K to X is a finite sequence S of sentences ending with X such that each sentence in S is either an assumption, deduced from a sentence (or more) in K, or deduced from previous sentences in S in accordance with one of N’s inference rules. These inference rules are introduction and elimination rules, defined for each logical constant of our language M. An introduction rule for a logical constant λ permits the introduction of a λ-sentence into a proof by allowing us to derive a sentence in which the logical constant appears. An elimination rule for a logical constant λ sentence allows us to derive a sentence from a λ-sentence that has at least one less occurrence of λ. An introduction rule for a logical constant is useful for deriving a sentence that contains the constant, and the elimination rule is useful for deriving a sentence from another in which the constant appears.

Formal proofs are not only epistemologically significant for securing knowledge, but also the derivations making up formal proofs may serve as models of the informal deductive reasoning that we perform. Indeed, a primary value of a formal proof is that it can serve as a model of ordinary deductive reasoning that explains the force of such reasoning by representing the principles of inference required to get to X from K. This is nice for, after all, like Molière’s M. Jourdain, who spoke prose all his life without knowing it, we frequently reason all the time without being aware of the principles underlying such reasoning.

Gentzen, one of the first logicians to present a natural deductive system, makes clear that a primary motive for the construction of his system is to reflect as accurately as possible the actual logical reasoning involved in mathematical proofs. He writes the following.

> My starting point was this: The formalization of logical deduction especially as it has been developed by Frege, Russell, and Hilbert, is rather far removed from the forms of deduction used in practice in mathematical proofs…In contrast, I intended first to set up a formal system which comes as close as possible to actual reasoning. The result was a 'calculus of natural deduction'. ((1934), p. 68)

Natural deductive systems are distinguished from other deductive systems by their usefulness in modeling ordinary, informal deductive inferential practices. Paraphrasing Gentzen, we may say that if one is interested in seeing logical connections between sentences in the most natural way possible, then a natural deductive system is a good choice for defining the deductive consequence relation.

Deductive System N

In stating N's rules, we begin with the simpler inference rules and give a sample formal deduction of them in action. Then we turn to the inference rules that employ what we shall call sub-proofs. In the statement of the rules, we let 'P' and 'Q' be any sentences from our language M. We shall number each line of a formal deduction with a positive integer. We let 'k', 'l', 'm', 'n', 'o', and 'p' be any positive integers (k < l < m < n < o).

&-Intro

k. P		k. Q	
l. Q		l. P	
m. (P & Q)	&-Intro: k, l	m. (P & Q)	&-Intro: k, l

& -Elim

k. (P & Q)		k. (P & Q)	
l. P	& Elim: k	l. Q	& Elim: k

&-Intro allows us to derive a conjunction from both of its two parts (called conjuncts). According to the &-Elim rule we may derive a conjunct from a conjunction. To the right of the sentence derived using an inference rule is the justification. Steps in a proof are justified by identifying both the lines in the proof used and by citing the appropriate rule.

~ -Elim

k. ~~ P
l. P ~-Elim: k

The ~ -Elim rule allows us to drop double negations and infer what was subject to the two negations.

v-Intro

k. P
l. (P v Q) v-Intro: k

k. P
l. (Q v P) v -Intro: k

By v-Intro we may derive a disjunction from one of its parts (called disjuncts).

→ -Elim

k. (P→Q)
l. P
m. Q →-Elim: k, l

The →-Elim rule corresponds to the principle of inference called *modus ponens*: from a conditional and its antecedent one may infer the consequent.

Here's a sample deduction using the above inference rules. The formal deduction–the sequence of sentences 4-11—is associated with the following pair.

<{(Female(paige) & Female (kelly)), (Female(paige)→~~Sister(paige, kelly)), (Female(kelly)→~~ Sister(paige, shannon))}, ((Sister(paige, kelly) & Sister(paige, shannon)) v Male(evan))>

The first element is the set of basis sentences and the second element is the conclusion. We number the basis sentences and list them (beginning with 1) ahead of the deduction. The deduction ends with the conclusion.

1. (Female(paige) & Female (kelly))	Basis
2. (Female(paige)→~~ Sister(paige, kelly))	Basis
3. (Female(kelly)→~~ Sister(paige, shannon))	Basis
4. Female(paige)	&-Elim: 1
5. Female(kelly)	&-Elim: 1
6. ~~ Sister(paige,kelly)	→-Elim: 2,4

7. Sister(paige,kelly) ~-Elim: 6
8. ~~ Sister(paige, shannon) →-Elim: 3,5
9. Sister(paige, shannon) ~-Elim: 8
10. (Sister(paige, kelly) & Sister(paige, shannon)) &-Intro: 7, 9
11. ((Sister(paige, kelly) & Sister(paige, shannon)) v Male(evan)) v-Intro: 10

Again, the (somewhat crooked) column all the way to the right gives the explanations for each line of the proof. Assuming the adequacy of N, the formal deduction establishes that the following inference is correct.

(Female(paige) & Female (kelly))
(Female(paige)→~~ Sister(paige, kelly))
(Female(kelly)→~~ Sister(paige, shannon))
∴((Sister(paige, kelly) & Sister(paige, shannon)) v Male(evan))

Recall from Chapter 2 that the symbol '∴' means 'therefore'. The set of sentences above the line is the basis of the inference, and the sentence below is the conclusion of the inference. For convenience in building proofs, we expand M to include '⊥', which we use as a symbol for a contradiction (e.g., '(Female(beth) & ~Female(beth))').

⊥-Intro

k. P		k. ~P	
l. ~P		l. P	
m. ⊥	⊥- Intro: k, l	m. ⊥	⊥- Intro: k, l

⊥-Elim

k. ⊥
l. P ⊥-Elim: k

If we have derived a sentence and its negation we may derive ⊥ using ⊥-Intro. The ⊥-Elim rule represents the idea that any sentence P is deducible from a contradiction. So, from ⊥ we may derive any sentence P using ⊥-Elim. Here's a deduction using the two rules.

1. Parent(beth,evan) & ~Parent(beth, evan) Basis
2. Parent(beth, evan) &-Elim: 1
3. ~ Parent(beth, evan) &-Elim: 1
4. ⊥ ⊥-Intro: 2,3
5.Parent(beth, shannon) ⊥-Elim: 4

For convenience, we introduce a reiteration rule that allows us to repeat steps in a proof as needed.

k. P
l. P   Reit: k

We now turn to the rules for the sentential connectives that employ what we shall call sub-proofs. Consider the following inference.

1. ~ (Married(shannon, kelly) & OlderThan(shannon, kelly))
2. <u>Married(shannon, kelly)</u>
∴ ~ Olderthan(shannon, kelly)

Here is an informal deduction of the conclusion from the basis sentences. Suppose that 'Olderthan(shannon, kelly)' is true. Then, from this assumption and basis sentence 2 it follows that '((Shannon is married to Kelly) & (Shannon is taller than Kelly))' is true. But this contradicts the first basis sentence '~((Shannon is married to Kelly) & (Shannon is taller than Kelly))', which is true by hypothesis. Hence our initial supposition is false. We have derived that '~(Shannon is married to Kelly)' is true.

Such a proof is called *a reductio ad absurdum* proof (or *reductio* for short). *Reductio ad absurdum* is Latin for something like 'reduction to the absurd'. In order to model this proof in N we introduce the ~-Intro rule.

~ -Intro

  k. P   Assumption
  l. ⊥

m. ~P   ~ -Intro: k-l

The ~-Intro rule allows us to infer the negation of an assumption if we have derived a contradiction, symbolized by '⊥', from the assumption. The indented proof margin (k-l) signifies a sub-proof. A sub-proof will be distinguished by indentation and spacing above its first line and below its last. In a sub-proof the first line is always an assumption (and so requires no justification), which is cancelled when the sub-proof is ended and we are back out on a line that sits on a wider imaginary proof margin. The effect of this is that we can no longer appeal to any of the lines in the sub-proof to generate later lines on wider proof margins. No deduction ends in the middle of a sub-proof. Here is a formal analogue of the above informal *reductio*.

1.~ (Married(shannon, kelly) & OlderThan(shannon, kelly)) Basis
2. Married(shannon, kelly) Basis

  3. OlderThan(shannon, kelly) Assumption
  4. (Married(shannon, kelly) & OlderThan(shannon, kelly)) &-Intro: 2, 3
  5. ⊥ ⊥-Intro: 1, 4

6. ~ Olderthan(shannon, kelly) ~ Intro: 3-5

Again, we signify a sub-proof with the indented list of its lines, and extra spacing above its starting line (e.g., line 3) and below its finish line (e.g., line 5). An assumption, like a basis sentence, is a supposition we suppose true for the purposes of the deduction. The difference is that whereas a basis sentence may be used at any step in a proof, an assumption may only be used to make a step within the sub-proof it heads. At the end of the sub-proof, the assumption is discharged. We now look at more sub-proofs in action and introduce another of N's inference rules. Consider the following inference.

1. (Male(kelly) v Female(kelly))
2. (Male(kelly)→ ~Sister(kelly, paige))
3. (Female(kelly)→~Brother(kelly, evan))

∴(~Sister(kelly,paige) v ~Brother(kelly, evan))

Informal proof: by assumption '(Male(kelly) v Female(kelly))' is true, i.e., by assumption at least one of the disjuncts is true. Suppose that 'Male(kelly)' is true. Then by *modus ponens* we may derive that '~Sister(kelly, paige)' is true from this assumption and the basis sentence 2. Then '(~Sister(kelly, paige) v ~Brother(kelly,evan))' is true. Now suppose that 'Female(kelly)' is true. Then by *modus ponens* we may derive that '~Brother(kelly,evan)' is true from this assumption and the basis sentence 3. Then '(~Sister(kelly, paige) v ~Brother(kelly, evan))' is true. So in either case we have derived that '(~Sister(kelly, paige) v ~Brother(kelly, evan))' is true. Thus we have shown that this sentence is a deductive consequence of the basis sentences.

We model this proof in N using the v-Elim rule.

v-Elim

k. (P v Q)

  l. P Assumption
  m. R

  n. Q Assumption

o. R

p. R v-Elim: k, l-m, n-o

The v-Elim rule allows us to derive a sentence from a disjunction by deriving it from each disjunct, possibly using sentences on earlier lines whose margins are less indented. The following formal proof models the above informal one.

1. (Male(kelly) v Female(kelly))	Basis
2. (Male(kelly)→ ~Sister(kelly, paige))	Basis
3. (Female(kelly)→~Brother(kelly, evan))	Basis
4. Male(kelly)	Assumption
5. ~Sister(kelly, paige)	→-Elim: 2, 4
6. (~Sister(kelly, paige) v ~Brother(kelly, evan))	v-Intro: 5
7. Female(kelly)	Assumption
8. ~Brother(kelly, evan)	→-Elim: 3,
9. (~Sister(kelly, paige) v ~Brother(kelly, evan))	v-Intro: 8
10. (~Sister(kelly, paige) v ~Brother(kelly, evan))	v-Elim: 1, 4-6,7-9

Here is N's representation of the principle of the disjunctive syllogism: for any sentences P and Q, from (P v Q) and ~P to infer Q.

1. P v Q	Basis
2. ~P	Basis
3. P	Assumption
4. ⊥	⊥-Intro: 2,3
5. Q	⊥-Elim: 4
6. Q	Assumption
7. Q	Reit: 6
8. Q	v-Elim: 1, 3-5, 6-7

Now we introduce the →-Intro rule by considering the following inference.

1. (Olderthan(shannon, kelly)→OlderThan(shannon, paige))
2. (OlderThan(shannon, paige)→OlderThan(shannon, evan))
∴ (Olderthan(shannon, kelly)→OlderThan(shannon, evan))

Informal proof: suppose that OlderThan(shannon, kelly). From this assumption and basis sentence 1 we may derive, by *modus ponens*, that OlderThan(shannon, paige). From this and basis sentence 2 we get, again by *modus ponens*, that OlderThan(shannon,evan). Hence,if OlderThan(shannon, kelly), then OlderThan(shannon,evan).

The structure of this proof is that of a conditional proof: a deduction of a conditional from a set of basis sentence which starts with the assumption of the antecedent, then a derivation of the consequent, and concludes with the conditional. To build conditional proofs in N, we rely on the →-Intro rule.

→-Intro

k. P Assumption
l. Q

m. (P→ Q) →-Intro: k-l

According to the →-Intro rule we may derive a conditional if we derive the consequent Q from the assumption of the antecedent P, and, perhaps, other sentences occurring earlier in the proof on wider proof margins. Again, such a proof is called a conditional proof. We model the above informal conditional proof in N as follows.

1. (Olderthan(shannon, kelly) → OlderThan(shannon, paige)) Basis
2. (OlderThan(shannon, paige) → OlderThan(shannon, evan)) Basis

3. OlderThan(shannon, kelly) Assumption
4. OlderThan(shannon, paige) →-Elim: 1,3
5. OlderThan(shannon, evan) →-Elim: 2,4

6. (Olderthan(shannon, kelly) → Olderthan(shannon, evan)) →-Intro: 3-5

Mastery of a deductive system facilitates the discovery of proof pathways in hard cases and increases one's efficiency in communicating proofs to others and explaining why a sentence is a logical consequence of others. For example, (1) if Beth is not Paige's parent, then it is false that if Beth is a parent of Shannon, Shannon and Paige are sisters. Suppose (2) that Beth is not Shannon's parent. Then we may infer that Shannon and Paige are sisters from (1) and (2). Of course, knowing the type of sentences involved is helpful for then we have a clearer idea of the inference principles that may be involved in deducing that Beth is a parent of Paige. Accordingly, we represent the two basis sentences and the conclusion in M, and then give a formal proof of the latter from the former.

1. (~ Parent(beth, paige) → ~ (Parent(beth, shannon) → Sister(shannon, paige))) Basis
2. ~ Parent(beth, shannon) Basis
   3. ~ Parent(beth, paige) Assumption
   4. ~ (Parent(beth, shannon) → Sister(shannon, paige)) → -Elim: 2, 3
      5. Parent(beth, shannon) Assumption
      6. ⊥ ⊥ -Intro: 3, 5
      7. Sister(shannon, paige) ⊥ -Elim: 6
   8. (Parent(beth, shannon) → Sister(shannon, paige)) → -Intro: 5-7
   9. ⊥ ⊥ -Intro: 4, 8
10. ~~ Parent(beth, shannon) ~ -Intro: 3-9
11. ⊥ ⊥ -Intro: 2, 10
12. Sister(shannon, paige) ⊥ -Elim: 11

Because we derived a contradiction at line 11, we got 'Sister(shannon, paige)' at line 12, using ⊥ -Elim. Look at the conditional proof (lines 5-7) from which we derived line 8. Lines 4 and 8 generated the contradiction. This is our first example of a sub-proof (5-7) embedded in another sub-proof (3-9). It is unlikely that independent of the resources of a deductive system, a reasoner would be able to readily build the informal analogue of this pathway from the basis sentences 1 and 2 to sentence 12. Again, mastery of a deductive system such as N can increase the efficiency of our performances of rigorous reasoning and cultivate skill at producing elegant proofs (proofs that take the least number of steps to get from the basis to the conclusion).

We now introduce the intro and elim rules for the identity symbol and the quantifiers. Let 'n' and 'n'' be any names, and 'Ωn' and 'Ωn'' be any well-formed formula in which n and n' appear and that have no free variables.

= -Intro

k. n=n = -Intro

= -Elim

k. Ωn

l. n=n'
m. Ωn' = -Elim: k. 1

The =-Intro rule allows us to introduce n=n at any step in a proof. Since n=n is deducible from any sentence, there is no need to identify the lines from which line k is derived. In effect, the =-Intro rule confirms that '(paige=paige)', '(shannon=shannon)', '(kelly=kelly)', etc… may be inferred from any sentence(s). The =-Elim rule tells us that if we have proven Ωn and n=n', then we may derive Ωn' which is gotten from Ωn by replacing n with n' in some but possibly not all occurrences. The =-Elim rule represents the principle known as the indiscernibility of identicals, which says that if n=n', then whatever is true of n is true of n'. This principle grounds the following inference.

1. ~ Sister(beth, kelly)
2. (beth=shannon)
∴~ Sister(shannon, kelly)

The indiscernibility of identicals is fairly obvious. If I know that Beth isn't Kelly's sister and that Beth is Shannon (perhaps 'Shannon' is an alias) then this establishes, with the help of the indiscernibility of identicals, that Shannon isn't Kelly's sister. Now we turn to the quantifier rules.

Let 'Ωv' be a formula in which 'v' is the only free variable, and let 'n' be any name.

∃-Intro

k. Ωn
l. ∃vΩv ∃-Intro: k

∃-Elim

k. ∃vΩv
[n] l. Ωn Assumption
m. P
n. P ∃-Elim: k, l-m

In the statement of the ∃-Elim rule, 'n' must be unique to the subproof, i.e., 'n' doesn't occur on any of the lines above l and below m. The ∃-Intro rule, which represents the principle of inference known as existential generalization, tells us that if we have proven Ωn, then we may derive ∃vΩv which results from Ωn by replacing n with a variable v in some but possibly not all of its occurrences and prefixing the existential quantifier. According to this rule, we may infer, say, '∃xMarried(x,matt)' from the sentence 'Married(beth, matt)'. By the ∃-Elim rule, we may reason from a sentence that is produced from an existential quantification by stripping the quantifier and replacing the resulting free variable in all of its occurrences by a name which is new to the proof. Recall that the language M has an infinite number of constants, and the name introduced by the ∃-Elim rule may be one of the them. We regard the assumption at line l, which starts the embedded sub-proof, as saying "Suppose *n* is an arbitrary individual from the domain of discourse such that Ωn."

To illustrate the basic idea behind the ∃-Elim rule, if I tell you that Shannon admires some McKeon, you can't infer that Shannon admires any particular McKeon such as Matt, Beth, Shannon, Kelly, Paige, or Evan. Nevertheless we have it that she admires somebody. The principle of inference corresponding to the ∃-Elim rule, called existential instantiation, allows us to assign this 'somebody' an arbitrary name new to the proof, say, 'w_1' and reason within the relevant sub-proof from 'Shannon admires w_1'. Then we cancel the assumption and infer a sentence that doesn't make any claims about w_1.

For example, suppose that (1) Shannon admires some McKeon. Let's call this McKeon 'w_1', i.e., assume (2) that Shannon admires a McKeon named 'w_1'. By the principle of inference corresponding to v-intro we may derive (3) that Shannon admires w_1 or w_1 admires Kelly. From (3), we may infer by existential generalization (4) that for some McKeon x, Shannon admires x or x admires Kelly. We now cancel the assumption (i.e., cancel (2)) by concluding using existential instantiation (5) that for some McKeon x, Shannon admires x or x admires Kelly from (1) and the subproof (2)-(4). Here is the above reasoning set out formally.

1. ∃xAdmires(shannon, x)	Basis
[w_1] 2. Admires(shannon, w_1)	Assumption
3. Admires(shannon, w_1) v Admires(w_1, kelly)	v-Intro: 2
4.∃x (Admires(shannon, x) v Admires(x, kelly))	∃-Intro: 3
5. ∃x (Admires(shannon, x) v Admires(x, kelly))	∃-Elim: 1, 2-4

The string at the assumption of the sub-proof (line 2) says "Suppose that w_1 is an arbitrary McKeon such that Admires(shannon, w_1)." This is not a sentence of M, but of the meta-language for M, i.e., the language used to talk about M. Hence, the ∃-Elim rule (as well as the ∀-Intro rule introduced below) has a meta-linguistic character.

∀-Intro

[n] k. Assumption
l. Ωn

m. ∀v Ωv ∀-Intro: k-l

∀-Elim

k. ∀vΩv
l. Ωn ∀-Elim: k

We add to the statement of the ∀-Intro rule that 'n' must be unique to the subproof. The ∀-Elim rule corresponds to the principle of inference known as universal instantiation: to infer that something holds for an individual of the domain if it holds for the entire domain. The ∀-Intro rule allows us to derive a claim that holds for the entire domain of discourse from a proof that the claim holds for an arbitrary selected individual from the domain. The assumption at line k reads in English "Suppose *n* names an arbitrarily selected individual from the domain of discourse." As with the ∃-Elim rule, the name introduced by the ∀-Intro rule may be one of the w_i consistent with the proviso that starts this paragraph. The ∀-Intro rule corresponds to the principle of inference often called universal generalization.

For example, suppose that we are told that (1) if a McKeon admires Paige, then that McKeon admires himself/herself, and that (2) every McKeon admires Paige. To show that we may correctly infer that every McKeon admires himself/herself we appeal to the principle of universal generalization, which (again) is represented in N by the ∀-Intro rule. We begin by assuming that (3) a McKeon is named 'w_1'. All we assume about w_1 is that w_1 is one of the McKeons. From (2), we infer that (4) w_1 admires Paige. We know from (1), using the principle of universal instantiation (the ∀-Elim rule in N), that (5) if w_1 loves Paige, then w_1 loves w_1. From (4) and (5) we may infer that (6) w_1 loves w_1 by *modus ponens*. Since w_1 is an arbitrarily selected individual (and so what holds for w_1 holds for all McKeons) we may conclude from (3)-(6) that (7) every McKeon loves himself/herself follows from (1) and (2) by universal generalization. This reasoning is represented by the following formal proof.

1. ∀x(Admires(x, paige)→Admires(x, x)) Basis
2. ∀x Admires(x, paige) Basis

[w_1] 3. Assumption
4. Admires(w_1, paige) ∀-Elim: 2
5. (Admires(w_1, paige) → Admires(w_1, w_1)) ∀-Elim: 1
6. Admires(w_1, w_1) →-Elim: 4,5

7. ∀x Admires(x, x) ∀-Intro: 3-6

Line 3, the assumption of the sub-proof, corresponds to the English sentence "Let w_1 be an arbitrary McKeon." The notion of a name referring to an arbitrary individual from the domain of discourse, which is utilized by both the ∀-Intro and ∃-Elim rules in the assumptions that start the respective sub-proofs, incorporates two distinct ideas. One, relevant to the ∃-Elim rule, means "some specific object, but I don't know which", while the other, relevant to the ∀-Intro rule means "any object, it doesn't matter which" (See Pelletier (1999), p. 118-120 for discussion). Consider the following set K of sentences and sentence X.

K={All McKeons admire those who admire somebody, Some McKeon admires a McKeon}
X= Paige admires Paige

Here's a proof that K produces X.

1. ∀x(∃y Admires(x, y) → ∀z Admires(z, x)) Basis
2. ∃x∃y Admires(x, y) Basis

[w_1] 3. ∃y Admires(w_1, y) Assumption
4. (∃y Admires(w_1, y) → ∀z Admires(z, w_1)) ∀-Elim:1
5. ∀z Admires(z, w_1) →-Elim: 3,4
6. Admires (paige, w_1) ∀- Elim: 5
7. ∃y Admires(paige, y) ∃ -Intro: 6
8. (∃yAdmires(paige, y) → ∀z Admires(z, paige)) ∀ -Elim: 1
9. ∀z Admires(z, paige) →-Elim: 7,8
10. Admires(paige, paige) ∀-Elim: 9

11. Admires(paige, paige) ∃-Elim: 2, 3-10

An informal correlate put somewhat succinctly, runs as follows. Let's call the unnamed admirer, mentioned in (2), w_1. From this and (1), every McKeon admires w_1 and so Paige admires w_1. Hence, Paige admires somebody. From this and (1) it follows that everybody admires Paige. So, Paige admires Paige. This is our desired conclusion. Even though the informal proof skips steps and doesn't mention by name the principles of inference used, the formal proof guides its construction.

The Deductive-Theoretic Characterization and the Common Concept of Logical Consequence

The case for the extensional adequacy of our deduction characterization of logical consequence may be put as follows.

(1) X is a logical consequence of K iff it is not logically possible for all the sentences in K to be true and X false, this is due to the forms of the sentences, and this is knowable *a priori*.

(2) It is not logically possible for all the sentences in K to be true and X false, this is due to the forms of the sentences, and this is knowable *a priori* iff X is deducible from K.

(3) X is deducible from K iff K $\vdash_N$ X.

Therefore, X is a logical consequence of K iff K $\vdash_N$ X.

We take the conclusion to assert the extensional adequacy of $\vdash_N$. Premise (1) summarizes Tarski's observation of the common concept of logical consequence, which was discussed in Chapter 2. There, it was highlighted that the logical consequence relation for a language L is determined by the meanings of L's logical constants, which are determinants of the logical forms of L-sentences. By the lights of the common concept of logical consequence, in order to fix what follows logically from what in a language we must select a class of constants that determines an *a priori*, formal consequence relation which reflects the modal element in the common concept of logical consequence. Hence, as discussed in Chapter 2, an account of how the meanings of logical constants determine the logical consequence relation will uniquely explain the necessity, formality, and *a prioricity* of the logical consequence relation.

In Chapter 2, we highlighted two different approaches to the meanings of logical constants: (1) in terms of the constant's truth-conditional properties, and (2) in terms of their inferential properties. We noted that each approach uniquely accounts for the necessity, formality, and *a prioricity* of the logical consequence relation, and yields a distinct characterization of logical consequence.

(a) X is a logical consequence of K iff there is no possible interpretation of the non-logical terminology of the language according to which all the sentence in K are true and X is false.

(b) X is a logical consequence of K iff X is deducible from K.

In Chapter 4 we made the notion of a *possible interpretation* in (a) precise by appealing to the technical notion of a *model*. Here in Chapter 5 we made the notion of *deducibility* in (b) precise by appealing to the technical notion of a *deductive system*. These attempts at precision yield the following theoretical characterizations of logical consequence.

(a′) The *model-theoretic characterization of logical consequence*: X is a logical consequence of K iff all models of K are models of X.

(b′) The *deductive- (or proof-) theoretic characterization of logical consequence*: X is a logical consequence of K iff there is a deduction in a deductive system of X from K.

To make it official, we instantiate the second characterization and characterize logical consequence in terms of deducibility in N.

(b″) A sentence X of M is a logical consequence of a set K of sentences from M iff X is deducible in N from K.

Statement (b″) is the conclusion of the above argument for the extensional adequacy of $\vdash_N$. The $\models$ and $\vdash_N$ relations are extensionally equivalent. That is, for any set K of M sentences and M-sentence X, K $\vdash_N$ X iff K $\models$X. A soundness proof establishes K $\vdash_N$ X only if K $\models$X, and a completeness proof establishes K $\vdash_N$ X if K $\models$X. With these two proofs in hand, we say that our deductive system N is complete and sound with respect to the model-theoretic consequence relation. Both the status of completeness and soundness proofs, and whether we regard premise (2) as an extensional or intensional equivalence turns on whether or not one characterization of the logical consequence relation is more basic than the other. This, in turn, depends on whether or not we identify the meanings of logical expressions with either their truth-conditional or inferential properties.

In this book, we regard the meaning of a logical term as a determinant of its truth-conditional and inferential properties, without identifying its meaning with either. If the meanings of logical terms are identified with their truth-conditional properties, then this motivates treating the model-theoretic characterization as basic and, if correct, taking the soundness of a proof-theoretic characterization to establish its correctness and its completeness to establish its utility in fixing what follows logically from what. Here premise (2) is an extensional equivalence. This raises the question of how the account of the truth-conditional properties of logical terms is justified. In Chapter 4, we said that what the truth-conditional properties of the logical terms are turns on the correct semantics of the non-logical terminology such as individual constants and predicates, and we sketched some reasons for being skeptical regarding the classical semantical treatment of the later.

If we identify the meanings of logical expressions with their inferential properties, then the proof-theoretic characterization is basic, and, if correct, the completeness of the proof-theoretic characterization establishes the correctness of the model-theoretic characterization and soundness establishes its utility. On this approach, premise (2) is an intensional equivalence, i.e., it explains the essential features of the concept of logical consequence in terms of deduction. This raises the question of how the account of the inferential properties of logical terms is justified. How do we justify rules of inference? Towards answering these questions, we consider a view we call "Inferentialism".

Inferentialism, Harmony, and the Justification of Inference Rules

Inferentialism is the view that the meaning of a logical constant is constituted by the inferential rules that govern its use. Inferentialism grounds the primacy of the characterization of logical consequence in terms of deducibility over the characterization in terms of possible interpretations . According to inferentialism, the pair of introduction and elimination rules for a logical expression λ *fix* its meaning. To say that an inference is valid is to say that its conclusion is a logical consequence of the initial premise(s) of the inference. There are inferences whose validity arises solely from the meanings of certain expressions occurring in them. Following Prior (1960), we say such inferences are analytically valid. To illustrate, consider the following valid inference.

<u>(Sister(paige, kelly) & Sister(paige, shannon))</u>
∴ Sister(paige, kelly)

Since this inference is justified by &-elim and since according to inferentialism &-elim along with the &-intro rules constitutes the meaning of '&', the validity of the inference arises solely from the meaning of '&'. Note that we may also eschew the appeal to inferentialism, and say that the inference is analytically valid because &-elim is derived from the meaning of '&' as specified by, say, a truth table.

Prior (1960) introduces a sentential connective, 'tonk', as a means of questioning the intelligibility of the concept of analytical validity.

"tonk"

k. P		k. (P tonk Q)	
l. (P tonk Q)	tonk-intro: k	l. Q	tonk-elim: k

Consider the following deduction, which makes use of the tonk-rules.

1. Sister(paige, kelly)
2. (Sister(paige, kelly) tonk Sister(evan, shannon)) tonk-intro: 1
3. Sister(evan, shannon) tonk-elim: 2

Since each step in the derivation is justified by a tonk rule, and since validity is transitive, it follows that the inference from 1 to 3 is valid due to the meaning of 'tonk'. Sentence (1) is true, because Paige is Kelly's sister. However, (3) is false; Shannon's brother is not her sister. Since a concept of validity that makes it non-truth-preserving is unintelligible, Prior's tonk-rule generates skepticism regarding the viability of analytic validity.

In essence, the above tonk-derivation demonstrates that the two inference-steps are permissible by the tonk rules, and yet not justified since they move us from truth to falsehood. (Stevenson (1961), p. 125) This highlights a potential defect in the conception of inferences whose validity arises solely from the meanings of certain expressions occurring in them. What is suggested is that "inference that is permissible by a meaning-determined rule" is not sufficient for "valid inference step". Hence, the notion of analytical validity is threatened. What's the moral that is to be drawn from Prior? We consider three responses to Prior's challenge to analytical validity.

One response takes Prior to have shown that the notion of analytical validity is unintelligible. According to this line, Prior has successfully stipulated a meaning for "tonk", and has shown that the tonk inference is permissible solely by virtue of the meaning of "tonk", but is nevertheless invalid. The connective is certainly useless: it's clearly evolutionary disadvantageous for a linguistic community to navigate reality using "tonk". Other responses aim to preserve the notion of analytical validity by arguing that Prior has failed to confer a meaning to the connective "tonk", because the proposed tonk-rules are illegitimate. One such response abandons inferentialism and claims that the tonk-rules are illegitimate, because they fail to reflect a meaning of "tonk" that is given independently of the tonk-rules and which constrains them. Another response holds that the tonk rules are defective, because they do not exhibit a harmony required of all such rules. The notion of harmony is then spelled out so as to preserve inferentialism. Let's itemize these three responses to Prior, and further discuss each.

The first response to Prior maintains that the fact that an inference is permissible by a meaning-determined rule is not sufficient for the inference to be valid. The inference must also be consistent with a plausible account of *deducibility*. We read '$\vdash$' as the sign for the deducibility relation, '$A_1, \ldots, A_n \vdash X$' says 'X is deducible from $A_1, \ldots, A_n$'. Clearly, we expect any deducibility relation to be truth-preserving: no falsehood is deducible from a truth. We follow Belnap (1961) and borrow from Gentzen's formulation of the intuitive properties of deducibility ((1935) pp.83-85) to give a theory of deducibility.

Axiom: $A \vdash A$

Rules:
Weakening: From $A_1, \ldots, A_n \vdash X$ infer $A_1, \ldots, A_n, B \vdash X$
Permutation: From $A_1, \ldots, A_i, A_{i+1}, \ldots, A_n \vdash X$ infer $A_1, \ldots, A_{i+1}, A_i, \ldots, A_n \vdash X$
Contraction: From $A_1, \ldots, A_n, A_n \vdash X$ infer $A_1, \ldots, A_n \vdash X$
Transitivity: From $A_1, \ldots, A_m \vdash X_1$ and $B_1, \ldots, B_n, X_1 \vdash X_2$ infer $A_1, \ldots, A_n, B_1, \ldots, B_n \vdash X_2$

In order for the inferences permitted by meaning-conferring inference principles to be valid, the deducibility relation such principles determine must be consistent with a correct account of the properties of *deducibility*. This induces a substantive constraint on which inference rules for logical constants determine *valid* inferences. Of course, one may say that if the deducibility relation determined by inference rules is not in sync with the correct account of deducibility, then those rules are not meaning-conferring. This is more in line with the third response below. In this book, we have followed Tarski and taken the meaning of logical constants to be a determinant of what follows logically from what. This commits us to maintaining the viability of analytical validity, and so this response is not available to us.

The following two responses to Prior reject that the tonk-rules define tonk. It is not the case that any rules that are stipulated for a connective will successfully fix that connective's meaning. The fact that the introduction and elimination rules for tonk permit an inference from truth to falsehood is evidence that the rules fail to confer meaning on the tonk-connective.

The second response to Prior rejects inferentialism, i.e., it is wrong to see the intro/elim rules as *defining* a logical expression λ. Rather, we must have an understanding of what λ means independently of its intro/elim rules in terms of which the latter are justified. Truth-tables are one way of spelling out this antecedent meaning. But what truth table justifies both tonk rules?

P	Q	(P tonk Q)
T	T	T
T	F	T
F	T	F
F	F	F

To make the intro-rule truth-preserving, we need to fill out the first two rows of the tonk-table as above. But then the elim-rule is clearly non-truth-preserving. The second row describes a case according to which P tonk Q is true and Q is false. Of course, any attempt to fix this by making P tonk Q false in the second row will invalidate the intro-rule. This suggests that 'tonk' is semantically inadequate since there is no unique truth-table that grounds both tonk rules. See Stevenson (1961) who adopts this response to Prior.

The third response to Prior maintains that the conditions that introduction and elimination rules must satisfy in order to confer meaning on the associated logical expressions can be spelled out so that inferentialism is preserved. The basic strategy for those who pursue this line of response to Prior is to claim that a pair of introduction and elimination rules are meaning conferring only if they exhibit a kind of harmony, and then to show that tonkish rules do not exhibit the required harmony. On this line of response to Prior, the harmony requirement has to be motivated so as to preserve inferentialism. There are three approaches to explaining the notion of harmony.

Harmony-as-conservativeness (Belnap), (Dummett)
Harmony-as-reduction (Prawitz), (Read)

Harmony-as-deductive equilibrium (Tennant)

The expression "harmony" was first introduced by Dummett ((1973), p. 396) in the course of discussing revision to ordinary linguistic practice. We only have room to discuss the first two approaches. See Tennant (2007) for the latest on the third approach. Before discussing each of the first two approaches, we develop the notion of harmony informally, working primarily off of Dummett and Prawitz. The primary aims of our discussion of harmony are to highlight justificatory criteria for the inference rules of a deductive system such as N, and clarify how exactly the introduction and elimination rules for a logical expression constitutes it meaning.

Harmony, informally considered

What, informally, is harmony? How, exactly, does it motivate the harmony requirement by means of which we preserve both the viability of analytical validity and inferentialism? The expression "harmony" was first introduced by Dummett (1973) in the course of discussing revision to ordinary linguistic practice, and then further developed in his (1991a).

> No one now supposes that established linguistic practice is sacrosanct. The supposition that it is, which was the fundamental part of 'ordinary language' philosophy, rested on the idea that, since meaning is use, we may adopt whatever linguistic practice we choose, and our choice will simply determine the use, and hence the meanings, of our expressions: any proposal to alter established practice is therefore merely a proposal to attach different senses to our expressions, whereas we have the right to confer on them whatever senses we please. The error underlying this attitude lies, in part, in the failure to appreciate the interplay between the different aspects of 'use', and the requirement of harmony between them. Crudely expressed, there are always two aspects of the use of a given form of sentence: the conditions under which an utterance of that sentence is appropriate, which include, in the case of an assertoric sentence, what counts as an acceptable ground for asserting it; and the consequences of an utterance of it, which comprise both what the speaker commits himself to by the utterance and the appropriate response on the part of the hearer, including, in the case of assertion, what he is entitled to infer from it if he accepts it. ((1973), p. 396)

> The two complementary features of [the conditions for the assertion of a sentence] ought to be in harmony with each other: and there is no automatic mechanism to ensure that they will be. The notion of harmony is difficult to make precise but intuitively compelling: it is obviously not possible for the two features of the use of any expression to be determined quite independently. Given what is conventionally

> accepted as serving to establish the truth of a given statement, the consequences of it as true cannot be fixed arbitrarily; conversely, given what accepting a statement as true is taken to involve, it cannot be arbitrarily determined what is to count as establishing it as true. ((1991a), p. 215)

We motivate the harmony requirement in a series of steps by elaborating on Dummett's above remarks. We draw heavily on both Dummett (1991a) and Prawitz (1977).

We start from the Wittgensteinian dictum: "The meaning of a sentence is exhaustively determined by its use." The next step sharpens the notion of use, relevant to motivating "harmony". Following Dummett ((1991, pp.210-211), we can distinguish two general categories of principles embodied in our linguistic practice. The first category consists of those that have to do with the circumstances that warrant an assertion (the *principles for verifying an assertion*). Such principles partly determine when we are entitled to make an assertion. The second category consists of those principles that determine the consequences that may be drawn from an utterance, i.e., a possible assertion, "the difference made by an utterance".

The distinction highlighted here is between two aspects of the assetoric use of a sentence (the use of a sentence to make an assertion).

(a) The conditions under which one is entitled to assert the sentence, i.e., the conditions under which a sentence can be correctly inferred.

(b) The consequences that one is entitled to draw from the sentence, i.e., the commitments incurred by asserting it.

Aspect (a) reflects a verificationist conception of meaning—the content of a statement is determined by what is required for its verification—and (b) reflects a pragmatic conception of meaning—the content of a statement is determined by its consequences for one who accepts it as true…the *difference made* by an utterance. (1991, p. 211)

The next step in our informal development of the notion of harmony elaborates on Dummett's remark: "The two features [of the assetoric use of a sentence] cannot be determined independently; on the contrary, each must be in full harmony with the other." The above summarizes the two ways of asserting a sentence: either directly by applying feature (a), or indirectly by inferring it from other sentences, and applying feature (b) to them (Prawitz (1977), p. 9). Prawitz remarks that,

> " …we have to require a harmony between these two ways of asserting a sentence. The indirect means of asserting a sentence must constitute only a conservative extension of the direct ones, i.e., they must not allow us to assert something which could not also be asserted by using direct means. ((1977), p. 9)

To illustrate, consider the following two ways of asserting a sentence A.

Situation (i): I assert A on the basis of circumstances which verify A.
Situation (ii): I assert A on the basis of inferring it from B.

There is disharmony if situation (ii) is allowable, while situation (i) is not because, say, there are no such circumstances which verify A. For example, "That is a dog" should not be assertible on the basis of inferring it from "That is a canine" if the circumstances that would directly verify "That is a dog" fail to obtain (say, because that, in fact, is a cat). In short, the requirement that the two aspects of the assertoric use of a sentence be in harmony is the requirement that "the principle for verifying an assertion should cohere with the consequences that can be drawn from it." (Read p. 126) In particular, "what follows from a statement should not outrun the grounds for its assertion." (Milne p. 56)

We instantiate this general notion of harmony by applying (a) and (b) to the assertibility conditions for λ-sentences where λ is a logical constant. Again, we use remarks made by Dummett as a guide.

> If we are wishing to formulate the meaning of, say, a binary sentential connective(), our task will be to explain the meaning of a sentence in which that connective is the principle operator, assuming the meanings of the two sub-sentences are given. If we are working in the context of a verificationist meaning-theory, we have to find a means of specifying what in general is to constitute a canonical means of verifying a statement made by uttering a sentence of the form A()B, given how A and B are to be verified. The hope is that this can be done by appeal to the introduction rules for the connective () in a natural deduction formalization of logic. ((1991a), p.215)
>
> If we are working in the context of a pragmatist meaning-theory, we shall have to explain the canonical means of drawing the consequences of a statement of the form A()B, given that we know the consequence of A and of B. The hope here will be that this can be done by appeal to the elimination rules for the connective in a natural deduction system. ((1991a), p.215)

Each logical expression λ is associated with a pair of introduction and elimination rules. The introduction rule for λ tells one how to infer a λ-sentence. The elimination rule tells one how to infer from a λ-sentence. The elimination rule permits one to eliminate the dominant occurrence of λ displayed in the premise (called the major premise of the elimination), and infer a conclusion which lacks a matching occurrence of λ. One way to look at these rules is that the introduction rule specifies the conditions under which one is permitted to assert a λ-sentence; the elimination rule specifies the consequences one is permitted to directly draw from a λ-sentence. In this context, the harmony requirement is the requirement that the introduction rule(s) for a logical constant λ, the rule(s) for inferring a λ-sentence, and the elimination rule(s) for λ, the rule(s) for drawing consequences from a λ-sentence, be harmonious. We draw from Prawitz in order to further clarify.

> Having a criterion for the acceptability of an inference, we have in particular a criterion for what commitments are made by an assertion, i.e., what consequences follow from the assertion of a sentence. Roughly speaking, we cannot infer other conclusions than such as must hold in view of the condition for asserting the sentence; we cannot take out more from the sentence than is put into it by these conditions. This is what is to be meant by saying that there must be a harmony between the rules for asserting a sentence and the rules for drawing consequences from it.

The introduction and elimination rules for a logical constant λ exhibit the required harmony only if given a canonical proof of a λ-sentence A from which we directly infer a sentence B *via* a λ-elim rule, B is deducible from the premises of the canonical proof of A. This reflects the idea that we cannot take more out of A than is put into it by the premises of its proof.

Harmony-as-conservativeness (Belnap) (Dummett)

Dummett writes that,

> Any one given logical constant, considered as governed by some set of logical laws, will satisfy the criterion for harmony provided that it is never possible, by appeal to those laws, to derive from premisses not containing that constant a conclusion not containing it and not attainable from those premisses by other laws that we accept. ((1991a), p.219)

Understanding "logical laws" in terms of inference principles, the motive for requiring that a constant's introduction and elimination rules be harmonious is grounded on, in Dummett's words, "our fundamental conception of what deductive inference accomplishes."

> An argument or proof convinces us because we construe it as showing that, given the premisses hold good according to our ordinary criteria, the conclusion must also hold *according to the criteria we already have for its holding* [italics are Dummett's].

In other words: deductive inference extends the criteria of assertibility from premises to conclusion, i.e., the standards for the assertibility of a conclusion correctly inferred from premises are those for the assertibility of the premises. Suppose that the rules for a constant λ exhibit disharmony because by their use we can take more out of a λ-sentence than is put into it by the premises of its proof. Then it is possible that what we conclude from the λ-sentence will not be assertible by the standards for the assertibility of its premises. If the later are assertible, then valid deductive inference would take us from assertible sentences to a conclusion that is not assertible by those standards. But this contradicts "our fundamental assumption of what deductive inference accomplishes". The harmony requirement is demanded by what we take deductive inference to accomplish. We now develop the above informal construal of harmony in terms of the notion of "conservativeness" paying close attention to Belnap (1962), where this notion is first used as a response to Prior.

Given a set R of intro/elim rules less those for a logical constant λ, "harmony" understood in terms of conservativeness is the relationship between the intro/elim rules for λ and those in R. For a language L, a set R of intro/elim rules for logical constants of L fixes the deducibility relation $\vdash_R$ defined on sentences of L. Specifically, where K is a set of L-sentences and X is an L-sentence, $K \vdash_R X$ (a deducibility statement relative to R) means that there is a proof of S from K using only intro/elim rules from R. If we extend L to L′ by adding a logical constant λ to L and extend R to R′ by adding the appropriate intro/elim rules for λ to R, then the resulting $\vdash_{R'}$-relation must be a conservative extension of $\vdash_R$. What this means is that for any set K of L′-sentences and L′-sentence X, if λ occurs neither in K nor in X, then $K \vdash_{R'} X$ only if $K \vdash_R X$. In other words: if $\vdash_{R'}$ is a conservative extension of $\vdash_R$, then there is no true deducibility statement, $K \vdash_{R'} X$ not containing λ, unless $K \vdash_R X$ is true.

Belnap's rationale for conservativeness is that "[t]he justification for unpacking the demand for consistency in terms of conservativeness is precisely our antecedent assumption that we already had *all* the universally valid deducibility-statements not involving any special connectives." ((1962) p.

132) From above, the set R of intro/elim rules for logical constants of L *fixes* the deducibility relation $\vdash_R$ defined on sentences of L. It is, therefore, unreasonable to think that the addition of a new connective can legitimately add to the extension of this relation.

Belnap maintains that we ought to also add *uniqueness* as a requirement for the introduction of connectives defined by inference rules. ((1962), p. 133) A constant λ is *unique* if for any constant λ′ whose intro/elim rules are formally identical with those of λ, it is the case that every formula containing λ is deductively equivalent with the formula containing λ′ in place of λ. The rationale for uniqueness is the inferentialist claim that a subset of the intro/elim rules for a constant λ fully determine its meaning. Clearly, this claim is threatened if there are two constants with formally identical intro/elim rules that bear different meanings.

Belnap draws the following moral from Prior.

> one *can* define connectives in terms of deducibility, but one bears the onus of proving at least consistency (existence); and if one wishes to further talk about *the* connective (instead of *a* connective) satisfying certain conditions, it is necessary to prove uniqueness as well. But it is not necessary to have an antecedent idea of the independent meaning of the connective. ((1962), p.134)

If we were to add tonk to our language M, and add its intro/elim rules to N generating N′, then clearly $\vdash_{N'}$ is not a conservative extension of $\vdash_N$. Hence, the intro/elim rules for tonk fail to define it. Hence, inferentialism is preserved and the threat Prior poses against the concept of analytical validity is deflated.

It turns out that N's ~-rules are not harmonious relative to the set R of N-inferential rules for v, &, and →. For example, $\vdash_R$ (((P→Q)→P)→P) is false, but true if we extend R to include N's ~-rules. This is evidence that the addition of N's ~-rules to R is not a conservative extension, and thus by the lights of this conception of harmony they fail to determine the meaning of '~'. Discussion of the significance of this follows the presentation of the next approach to harmony.

Harmony-as-reduction (Prawitz), (Read)

Gentzen makes the following remark on the relationship between the introduction and elimination rules in his natural deduction system.

> The introductions represent, as it were, the 'definitions' of the symbols concerned, and the eliminations are no more, in the final analysis, than the consequences of these definitions. This fact may be expressed as follows: In eliminating a symbol, we may use the formula with whose terminal symbol we are dealing only in the sense afforded it by the introduction of that symbol. ((1935), p. 80)

It is hard to see how the elimination rules are consequences of—in the sense of being deducible from—the introduction rules. At the very least, Gentzen means to say that an introduction rule in some sense justifies the corresponding elimination rule. He goes on to explain.

> An example may clarify what is meant: we were able to introduce the formula A→B when there existed a derivation of B from the assumption formula A. If we then wished to use that formula by eliminating the →-symbol (we could, of course, also use it to form longer formulae, e.g., (A→B) v C, v-Intro), we could do this precisely by inferring B directly once A has been proved, for what A→B attests is just the existence of a derivation of B from A. Note that in saying this we need not go into the 'informal sense' of the →-symbol. ((1935), pp.80-81)

Gentzen remarks that, "[b]y making these ideas more precise it should be possible to display the *E*-inferences as unique functions of their corresponding *I*-inferences on the basis of certain requirements." ((1935, p.81) Any proof ending with an application of an elimination rule can be reduced to a simpler proof if that elimination occurs immediately after a corresponding introduction.

Prawitz understands Gentzen's remarks as suggesting that the intro-rule(s) for a logical constant λ give(s) the meaning of 'λ' in terms of a canonical proof of a λ-sentence. The proof is described as "canonical" in order to distinguish it from other proofs that do not conclude a λ-sentence on the basis of an λ-intro rule. Prawitz observes that the elimination rule is, in a sense, the inverse of the corresponding introduction rule: by an application of an elimination rule one essentially only restores what had already been established if the major premise of the application was inferred by an application of an introduction rule. This relationship between the introduction rules and the elimination rules is roughly expressed by what Prawitz calls the *inversion principle*.

> "Let α be an application of an elimination rule that has B as consequence. Then, deductions that satisfy the sufficient condition ... for deriving the major premiss of α, when combined with deductions of the minor premisses of α (if any), already

"contain" a deduction of B; the deduction of B is thus obtainable directly form the given deductions without the addition of α." Prawitz (1965, p. 33)

Prawitz observes that the inversion principle says in effect that nothing is "gained" by inferring a formula through introduction for use as a major premise in an elimination.

Prawitz takes the inversion principle to formally express the harmony requirement. "The required harmony between the rules for inferring a sentence and the rules for drawing consequences from it is clearly related to the inversion principle for introduction and elimination rules in Gentzen's system of natural deduction." ((1977), p. 9) And later on, he remarks that. "[t]he inversion principle [...] is just an expression of the fact that the introduction rules and the elimination rules of Gentzen's system of natural deduction is in a harmony of this kind, and the normalization theorem which follows from this principle tells us that we can avoid all detours consisting of an introduction followed by an elimination." (Ibid., p. 25)

The normalization theorem for a deductive system S says that if $K \vdash_S X$, then there is a deduction in S of X from K in which no formula occurs both as a consequence of an application of an introduction rule and as a major premise of an application of an elimination rule ((1965), p.34). A *maximum formula* is both the conclusion of an application of an introduction rule and the major premise of an inference justified by an elimination rule. A derivation D is in normal form if it cannot be further reduced and so contains no maximum formula. Normalization is a procedure converting any derivation into normal form. Typically, normal form is achieved for a derivation through the application of reduction schemes for removal of maximum formulae in the derivation.

We now give reduction schemes for the elimination rules of our system N (we ignore $\perp$ and =). The reduction schemes illustrate how any proof ending with an application of an elimination rule can be reduced to a simpler proof if that elimination occurs immediately after the employment of a corresponding introduction rule. We loosely follow Prawitz, in adopting the following conventions.

D_i
A_i

This represents a derivation D_i that ends with A_i.

A_i A_i
D_i D_i
A_j

Either represents a derivation that depends on an assumption A_i. The open assumptions of a derivation are the assumptions on which the end formula depends. We say that a derivation is closed if it contains no open assumptions, and open if it does. For each of the below reduction schemes, the first proof scheme is reduced to the one that follows. The first proof scheme displays a derivation in which the last step is arrived at *via* an elimination rule. This derivation represent proof schemes canonical for the relevant elimination rule. Also, there occurs a maximum formula, which is derived canonically by the relevant introduction rule. For each reduction scheme, the second schematic proof has the same or fewer assumptions as the first proof, and lacks the detour through the maximum formula in the derivation of the conclusion. Instead of using line numbers in justifications, we display the relevant premises themselves.

The scheme of →-elim reduction

D_1
A_1
  A_1
  D_2
  A_2
$A_1 \to A_2$ →-Intro: A_1-A_2
A_2 →-Elim: A_1, $A_1 \to A_2$

This reduces to the following.

D_1
A_1
D_2
A_2

In essence, the two proof schemes illustrate how given a canonical derivation of $A_1 \to A_2$, a derivation of the antecedent A_1, and a derivation of the consequent A_2 from $A_1 \to A_2$ and A_1 using →-elim, we may construct the second proof which derives A_2 without the use of →-elim. Here is a simple instantiation of the reduction scheme. The first proof is reduced to the second one.

1. $(A_1 \& A_2)$
2. A_1 &-elim: 1
    3. A_1
    4. $(A_1 \vee A_2)$ v-intro: 3
5. $(A_1 \rightarrow (A_1 \vee A_2))$ →-intro: 3-4
6. $(A_1 \vee A_2)$ →-elim: 2,5

Reduction

1. $(A_1 \& A_2)$
2. A_1 &-elim: 1
3. $(A_1 \vee A_2)$ v-intro: 2

The scheme of &-elim reduction

D_1
A_1
D_2
A_2
$A_1 \& A_2$ &-intro: A_1, A_2
$A_{i\ (i=1\ \text{or}\ i=2)}$ &-elim: $A_1 \& A_2$

Reduction

D_i
A_i

The scheme of v-elim reduction

D_i
A_i
$A_1 \vee A_2$ v-intro: $A_{i(i=1\ \text{or}\ i=2)}$

  A_1
  D_2
  A_3

  A_2
  D_3
  A_3

A_3 v-elim: A_1 v A_2, A_1-A_3, A_2-A_3

Reduction

D_i
A_i
$D_{k\ (k=1\ \text{or}\ k=2)}$
A_3

Here's an instantiation of the scheme of v-elim reduction.

1. A_1
2. A_2
3. A_2 v A_3 v-intro: 2

    4. A_2
    5. (A_1 & A_2) &-intro: 1,2
    6. ((A_1 & A_2) v (A_1 & A_3)) v-intro: 5

    7. A_3
    8. (A_1 & A_2) &-intro: 1,2
    9. ((A_1 & A_2) v (A_1 & A_3)) v-intro: 8

10. ((A_1 & A_2) v (A_1 & A_3)) v-elim: 3, 4-6, 7-9

Reduction

1. A_1
2. A_2
3. A_1 & A_2 &-intro: 1,2
4. ((A_1 & A_2) v (A_1 & A_3)) v-intro: 3

There is no reduction scheme for our system N's ~-elim rule. So, N's introduction and elimination rules for negation do not exhibit "harmony" as understood in terms of Prawitz's development of Gentzen. Below there is discussion of the significance of this.

The scheme of ∃-elim reduction

D_1
P(t)
∃x P(x) ∃-intro: P(t)

    [t′] P(t′)
        $D_{2\ [t']}$

A_1

A_1 ∃-elim: ∃x P(x), [t′]P(t′)-C

Reduction

D_1
P(t)
$D_{2[t]}$
A_1

$D_{2[t]}$ is a deduction derived from $D_{2[t']}$ by replacing individual constant ‘t′’ everywhere in the later with ‘t’.

The scheme of ∀-elim reduction

[t]
$D_{1[t]}$
P(t)

∀x P(x) ∀-intro: [t]-A(t)
P(t′) ∀-elim: ∀x A(x)

Reduction

$D_{1[t']}$
P(t′)

Here is an example of this reduction scheme.

[a] P(a)
P(a) Reit: [a]P(a)

∀x (P(x)→P(x)) ∀-intro: [a]P(a)-P(a)
P(a)→ P(a) ∀-elim: ∀x(P(x)→P(x))

Reduction

P(a)
P(a)

P(a)→P(a) →-intro: P(a)-P(a)

According to Prawitz, the introduction rule(s) for a logical constant λ give(s) the meaning of ‘λ’ in terms of a canonical proof of a λ-sentence. That is, the meaning of λ is specified by stating the conditions under which one may directly infer a conclusion with that operator dominant. The reduction schemes for the elimination rules justify these rules by demonstrating that by using them one cannot infer from canonically derived sentences any conclusion that cannot be derived without using the elimination rules. The reduction schemes guarantee that the elimination rules conservatively extend the introduction rules, because they guarantee that direct proofs of the premises of an elimination rule can be transformed into a direct proof of its conclusion. A logical expression's introduction and elimination rules exhibit harmony only if there is a reduction scheme for the elimination rules. Harmony, in the context of Gentzen's view that the elimination rules follow from the introduction rules, illustrates that no more is inferable from a λ-sentence than the λ-intro rules permit. Since an introduction rule is constitutive of the meaning of its logical operator, it is self-justificatory. The justification for elimination rules is derived from the justification of the introduction rules *via* the reduction schemes.

There is no tonk reduction scheme. That is, there is no procedure for reducing the below proof-scheme to one of A_2 from the same or fewer assumptions where the form of the proof depends on the form of the proof-scheme as illustrated above.

D
A_1
A_1 tonk A_2 tonk-intro: A_1
A_2 tonk-elim: A_1 tonk A_2

This means that the tonk-rules do not exhibit the required harmony. The tonk-elim rule is not justified by the tonk-intro rule.

Whereas "harmony" understood in terms of conservativeness deals with the relationship between the rules for a logical constant λ and an antecedently-given set of inference rules, "harmony" understood in terms of the inversion principle deals directly with the relationship between the introduction and elimination rules. As noted above, according to each conception of "harmony", N's negation rules do not exhibit harmony. If either of these conceptions is correct, then this shows that N's negation rules fail to determine the meaning of '~'. We now consider the significance of this.

A deducibility theory for a language L is a set of inference rules defined for L-sentences. Relative to a language L, two deducibility theories D, D′ are extensionally equivalent iff for any set K of L-sentences and L-sentence X, $K \vdash_{D} X$ iff $K \vdash_{D'} X$. A (partial) logic for a language is simply the collection of equivalent deducibility theories. We say that a set S of N's inference rules expresses a logic LOG for a first-order language L if S ∈ LOG. Naturally, different subsets of N's inference rules express different logics.

Minimal logic is expressed by the introduction and elimination rules for &, v, and →, and the introduction rules for ~ and ⊥. Intuitionistic logic is expressed by the previous rules, plus the ⊥-elim rule. Classical logic is expressed by the previous rules that express intuitionistic logic, plus the ~-elim rule. The set of N-rules for intuitionistic logic is a conservative extension of the set of N-rules for minimal logic. However, the set of N-rules that express classical logic is not a conservative extension of the rules that express intuitionistic logic. Thus, by the first conception of "harmony", the inference rules for classical logic do not exhibit harmony. Following Dummett, we might conclude that classical logic fails to capture the meaning of negation (understanding meaning in terms of use). Also, since the ~-elim rule lacks a reduction procedure, the rules for classical logic are inharmonious by the lights of the second conception of harmony. Following Prawitz, we might conclude that classical logic fails to capture the meaning of negation.

Is it correct to draw these conclusions? Clearly not unless the respective accounts of harmony are adequate. But even if either is adequate, from the fact that N's inference do not exhibit harmony it does not follow that classical logic fails to capture the meaning of negation. This is true only if relative to one construal of "harmony" every deducibility theory, not just N, that expresses classical logic fails to exhibit the harmony requirement.

This is not the case when harmony is understood in terms of conservativeness. There are deducibility theories such as Gentzen's sequent calculus which express classical logic and according to which no deduction from a negation-free set K of sentences to a negation-free sentence X appeals to the theory's rules for negation. Of course, N's ~-rules may nevertheless be seen as defective for reasons dealing with the ease of building proofs with them. But any defect here is not a strike against the classical logical understanding of negation.

There is substantive criticism of the account of harmony in terms of conservativeness (Read (1988), Chapter 9). Here we observe the following in order to point to why conservativeness might be neither necessary nor sufficient for harmony. Intuitively, if the set of allowable conclusions from applications of the λ-elimination rules to an λ-sentence does not include all that is deducible from the grounds in a canonical proof of that λ-sentence, then introduction and elimination rules for λ are not in harmony. However, as Dummett acknowledges, conservativeness does not rule out such a defect.

In order to account for this type of defect, Dummett extends his characterization of harmony in terms of conservativeness by appealing to Prawitz's notion of normalization. He calls the latter 'intrinsic harmony' and the former 'total harmony' ((1991, p. 250).

Prawitz demonstrates that conservativeness is not necessary for harmony. Add a truth predicate to the language of arithmetic whose introduction and elimination rules are as follows (in the sentence, *T(P)*, we let *P* be a name for itself).

P
T(P) T-intro: P

T(P)
P T-elim: T(P)

These rules are clearly harmonious. The conclusion reached *via* the application of the T-elim rule is derivable from the basis for inferring the premise *via* the introduction rule. But this addition to the language of arithmetic allows us to prove a sentence (the Gödel sentence) which is not provable without it and is thus a non-conservative extension.

Read derives a notion of harmony from Prawitz's inversion principle. For a logical connective δ,

> if $\{\Pi_i\}$ denotes the grounds for introducing some formula A (introducing an occurrence of a connective δ in A), then the elim rule for δ should permit inference to an arbitrary formula C only if $\{\Pi_i\}$ themselves entail C. (We need to permit multiple introduction-rules here, to accommodate, for example, the case of disjunction, 'v'. There may be a whole variety of grounds for asserting A.) ((2000), p. 130)

Adopting Read's style of display for inference rules, the schematic form of the introduction rule(s) for a connective δ is as follows.

$$\frac{\Pi_1}{A}\ \delta\text{-Intro} \qquad \frac{\Pi_2}{A}\ \delta\text{-intro} \qquad \ldots$$

From this the schematic form of the elimination rules is derived.

$$\frac{A \quad \begin{matrix}(\Pi_1)\\ C\end{matrix} \quad \begin{matrix}(\Pi_2)\\ C\end{matrix} \quad \cdots}{C}\ \delta\text{-elim}$$

This schematic display illustrates that in cases where an application of δ-elim immediately follows an application of δ-intro, it is possible to produce in the manner of Prawitz a shorter and more direct proof of C which eliminates the maximum formula A. That is, suppose that we have both the following.

$$\frac{\Pi_i}{A}\ \delta\text{-intro}$$

$$\frac{A \quad \overset{(\Pi_1)}{C} \quad \overset{(\Pi_2)}{C} \quad \cdots}{C}\ \delta\text{-elim}$$

Then we may produce the following more direct route to C.

$$\frac{\Pi_i}{C}$$

We now look at two examples. Here's how the rules for '&' look when they instantiate the above schematic displays of the forms of harmonious introduction and elimination rules. (Read (2004), p. 114-115)

$$\frac{p \quad q}{p \,\&\, q}\ \&I$$

$$\frac{p \,\&\, q \quad \overset{(p,q)}{r}}{r}\ \&E'$$

The general form of the elim-rule for '&', &E′, simplifies to rules akin to our &-elim.

$$\frac{p \,\&\, q}{p}\ \&E$$

$$\frac{p \,\&\, q}{q}\ \&E$$

Read writes that "The idea behind the simplification is that "rather than infer from p & q what we can infer from p and q, we can simply infer p and q and then proceed what we can infer from each. ((2004), p.114)

Here are the introduction and elimination rules for '→'.

$$\frac{\overset{(p)}{q}}{p \rightarrow q}\ \rightarrow I$$

From a proof of q from an assumption p, we may discharge the assumption and infer p→q.

$$\frac{p\to q \quad \begin{matrix}(p)\\ q\\ r\end{matrix}}{r}\ \to E'$$

As Read demonstrates, this general form of the elimination rule for '→' is amenable to simplification. If we may infer whatever we can infer from a derivation of q from p, we can simply infer q from p and then proceed to infer whatever it is that we can infer from that inference. This is portrayed as follows.

$$\frac{p\to q \quad p}{q}\ \to E$$

This is akin to our system N's →-elim rule.

On Read's approach to harmony, Prior's tonk-rules fail to exhibit harmony because they do not instantiate the above general forms of introduction and elimination rules. If they were to instantiate such forms, they would look like the following.

$$\frac{p}{p \text{ tonk } q}\ \text{tonk- I}$$

$$\frac{p \text{ tonk } q \quad \begin{matrix}(p)\\ r\end{matrix}}{r}\ \text{tonk-E}'$$

As with the general forms of the elimination rule for '&' and '→', tonk-E′ is amenable to simplification. If we may infer whatever r we can infer from a derivation of p tonk q from p, we can simply infer p from p tonk q and then proceed to infer whatever it is that we can infer from that inference. The simplified version of the elimination rule for tonk is the following.

$$\frac{p \text{ tonk } q}{p}\ \text{tonk-E}$$

So, on Read's account of harmony Prior misstated the elimination rule for tonk. We now conclude with four observations derived from Read.

First, according to Read an expression is logical iff there exists harmonious introduction and elimination rules for it. ((2004), p. 115) The identity symbol is typically regarded as logical, and so by this criterion for logical constancy it should have harmonious rules. N's = -intro rule is too weak to justify = -elim, and so in terms of Gentzen's conception of harmony these rules are not in harmony. Read formulates an inference rule that reflects the identity of indiscernibles ($a{=}b =_{df}$ for every predicate F, F(a) iff F(b)), and uses it as the introduction rule for '='. Then he show that this rule justifies a version of N's elimination rule. The proposal motivates replacing N's =-elim rule with Read's =-intro rule. The details are in Read's (2004).

Second, there are deducibility theories for modal logic whose rules exhibit harmony in Read's sense, but for which normalization fails. This motivates skepticism regarding Prawitz's view that normalization is a consequence of harmony, informally considered. Although both Read's and Prawitz's conceptions of harmony are developments of Gentzen's undeveloped view, they are competitors.

Third, that the rules for a logical expression are harmonious in Read's sense is not sufficient for them to successfully determine the meaning of the logical expression. To illustrate this we appeal to Read's zero-place connective '•' whose rules are as follows. 'C' denotes the conclusion of the elim-rule inference.

$$\frac{\sim\bullet}{\bullet}\ \bullet\text{-I}$$

$$\frac{\bullet \quad \begin{matrix}(\sim\bullet)\\ C\end{matrix}}{C}\ \bullet\text{-E}$$

The rules for '•' exhibit harmony, yet when added to system N (after dressing N's rules in terms of Read's generalized forms for introduction and elimination rules) permits a derivation of 'Male(beth) & ~Male(beth)' from the truth '~Male(beth)'.

Finally, there are many deducibility theories for classical logic and many more when we consider other logics. This broad range of inference rules complicates both the explication and defense of harmony understood in terms of Gentzen's view. For example, in Tennant's deducibility theory for minimal logic (1978), N's ⊥-intro rule is Tennant's ~-elim rule, which is justified by a rule akin to our ~-intro rule. *Prima facie*, this suggests (i) that an introduction rule may be in harmony with two distinct elimination rules, and, therefore, uniquely justify neither. Furthermore, there is a deducibility theory for relevant logic whose →-intro rule is different from N's, yet it justifies an →-elim rule akin to N's. (See below, and for discussion see Hjortland Thomassen (2007)). It appears (ii) that one elimination rule may be harmoniously matched with different introduction rules. (For an example drawn from modal logic and further discussion see Read (2008)) Both (i) and (ii) make it hard to see how a logical constant's set of elimination rules is justified by its meaning as this is constituted by the set of its introduction rules. See the previous two references for further discussion.

Tarski's Criticism of the Deductive-Theoretic Characterization of Logical Consequence

Tarski (1936) claims that no deductive-theoretic characterization of logical consequence is extensionally equivalent with the common concept of logical consequence. ((1936), p. 411) On this reading, Tarski rejects that for every set K of sentences and sentence X, K $\vdash_N$ X iff it is not logically possible for all the sentences in K to be true and X false, this is due to the forms of the sentences, and this is knowable *a priori*. Tarski is critical of premise (3) in the above argument for the extensional adequacy of $\vdash_N$. Here's a rendition of his reasoning, focusing on $\vdash_N$ defined on a language for arithmetic, which allows us to talk about the natural numbers 0, 1, 2, 3, and so on. Let 'P' be a predicate defined over the domain of natural numbers and let 'NatNum (x)' abbreviate 'x is a natural number'. According to Tarski, intuitively,

$$\forall x(\mathrm{NatNum}(x) \rightarrow P(x))$$

is a logical consequence of the infinite set *S* of sentences, P(0), P(1), P(2), and so on. However, the universal quantification is not a $\vdash_N$ –consequence of the set S. The reason why is that the $\vdash_N$ –consequence relation is compact: for any sentence X and set K of sentences, if X is a $\vdash_N$ –consequence of K, then X is a $\vdash_N$ –consequence of some finite subset of K. Proofs in N are objects of finite length; recall that a deduction in N is a finite sequence of sentences. Since the universal quantification is not a $\vdash_N$ –consequence of any finite subset of S, it is not a $\vdash_N$–consequence of S. By the completeness of system N, it follows that

$$\forall x(\mathrm{NatNum}(x) \rightarrow P(x))$$

is not a $\models$-consequence of S either. Consider the structure **U** whose domain is the set of McKeons. Let all numerals name Beth. Let the extension of

'NatNum' be the entire domain, and the extension of 'P' be just Beth. Then each element of S is true in **U,** but '∀x(NatNum(x)→P(x))' is not true in **U**. Note that the sentences in S only say that P holds for 0, 1, 2, and so on, and not also that 0,1, 2, etc… are all the elements of the domain of discourse. The above interpretation takes advantage of this fact by reinterpreting all numerals as names for Beth.

However, we can reflect model-theoretically the intuition that '∀x(NatNum(x)→P(x))' is a logical consequence of set S by doing one of two things. We can add to S the functional equivalent of the claim that 1, 2, 3, etc., are all the natural numbers there are on the basis that this is an implicit assumption of the view that the universal quantification follows from S. Or we could add 'NatNum' and all numerals to our list of logical terms. On either option it still won't be the case that '∀x(NatNum(x)→P(x))' is a $\vdash_N$–consequence of the set S. There is no way to accommodate the intuition that '∀x(NatNum(x)→P(x))' is a logical consequence of S in terms of a compact consequence relation. Tarski takes this to be a reason to favor his model-theoretic account of logical consequence, which, as indicated above, can reflect the above intuition, over any compact consequence relation such as $\vdash_N$.

This forces us to grant that the notion of deducibility is not completely captured by a compact consequence relation such as $\vdash_N$, and acknowledge that premise (3) in the argument for the extensional adequacy of N is false. It needs to be weakened to:

X is deducible from K if K $\vdash_N$ X.

Correspondingly, the conclusion needs to be weakened to:

X is a logical consequence of K if K $\vdash_N$ X.

To be clear, the considerations here do not bring premise (2) into question. The idealized conception of deducibility in terms of a compact consequence relation is successful only with respect to a deductive consequence relation that turns on the inferential properties of some terms such as the sentential connectives and quantifiers of our language M. What Tarski's illustration shows is that what is called the ω-rule is a correct inference rule. The ω-rule is: <P(0), P(1), P(2), …> to infer ∀x(NatNum(x)→P(x)), with respect to any predicate P. Any inference guided by this rule is correct even though it can't be represented in a deductive system as this notion has been construed here.

We now sketch considerations that question the truth of the weakened conclusion.

Is N a Correct Deductive System?

Are there deductions in N that are intuitively incorrect? In order to fine-tune the question note that the sentential connectives, the identity symbol, and the quantifiers of M are intended to correspond to *or*, *and*, *not*, *if…then* (the indicative conditional), *is identical with*, *some*, and *all*. Hence, N is a correct deductive system only if the introduction and elimination rules of N reflect the inferential properties of the ordinary language expressions. Is this the case? Here we sketch three views critical of the correctness of system N.

Not everybody accepts it as a fact that any sentence is deducible from a contradiction, and so some question the correctness of the ⊥-Elim rule. Consider the following informal proof of Q from P &~P, for sentences P and Q, as a rationale for the ⊥-Elim rule.

From (1) P and not-P we may correctly infer (2) P, from which it is correct to infer (3) P or Q. We derive (4) not-P from (1). (5) P follows from (3) and (4).

The proof seems to be composed of valid modes of inference. Critics of the ⊥-Elim rule are obliged to tell us where it goes wrong. Here we follow the relevance logicians Anderson and Belnap (1962) pp.105-108 (for discussion, see Read (1995) pp. 54-60). In a nutshell, Anderson and Belnap claim that the proof is defective because it commits a fallacy of equivocation. The move from (2) to (3) is correct only if *or* has the sense of *at least one*. For example, from *Kelly is female* it is legit to infer that at least one of the two sentences *Kelly is female* and *Kelly is older than Paige* is true. On this sense of *or* given that Kelly is female, one may infer that Kelly is female or whatever you like. However, in order for the passage from (3) and (4) to (5) to be legitimate the sense of *or* in (3) is *if not-…then*. For example from *if Kelly is not female, then Kelly is not Paige's sister* and *Kelly is not female* it is correct to infer *Kelly is not Paige's sister*. Hence, the above "support" for the ⊥-Elim rule is defective for it equivocates on the meaning of *or*.

Two things to highlight. First, Anderson and Belnap think that the inference from (2) to (3) on the *if not-…then* reading of *or* is incorrect. Given that Kelly is female it is problematic to deduce that if she is not then Kelly is older than Paige—or whatever you like. Such an inference commits a fallacy of relevance for Kelly not being female is not relevant to her being older than Paige. The representation of this inference in system N appeals to the ⊥-Elim rule, which is rejected by Anderson and Belnap. Second, the principle of inference underlying the move from (3) and (4) to (5)—from P or Q and not-P to infer Q—is called the principle of the disjunctive syllogism. Anderson and Belnap claim that this principle is not generally valid when *or* has the sense of *at least one*, which it has when it is rendered by 'v' understood by the lights of classical semantics. If Q is relevant to P, then the principle holds on this reading of *or*.

It is worthwhile to note the essentially informal nature of the debate. It calls upon our pre-theoretic intuitions about correct inference. It would be quite useless to cite the proof in N of the validity of disjunctive syllogism (p. 51) against Anderson and Belnap for it relies on the ⊥-Elim rule whose legitimacy is in question. No doubt, pre-theoretical notions and original intuitions must be refined and shaped somewhat by theory. Our pre-theoretic notion of correct deductive reasoning in ordinary language is not completely determinant and precise independently of the resources of a full or partial logic (see Shapiro (1991) Chapters 1 and 2 for discussion of the interplay between theory and pre-theoretic notions and intuitions). Nevertheless, hardcore intuitions regarding correct deductive reasoning do seem to drive the debate over the legitimacy of deductive systems such as N and over the legitimacy of the ⊥-Elim rule in particular. Anderson and Belnap write that denying the principle of the disjunctive syllogism, regarded as a valid mode of inference since Aristotle, "…. will seem hopelessly naïve to those logicians whose logical intuitions have been numbed through hearing and repeating the logicians fairy tales of the past half century, and hence stand in need of further support" (p. 108). The possibility that intuitions in support of the general validity of the principle of the disjunctive syllogism have been shaped by a bad theory of inference is motive enough to consider argumentative support for the principle and to investigate deductive systems for relevance logic.

A natural deductive system for relevant logic has the means for tracking the relevance quotient of the steps used in a proof and allows the application of an introduction rule in the step from A to B "only when A is relevant to B in the sense that A is *used* in arriving at B" (Anderson and Belnap 1962, p. 90). Consider the following proof in system N.

1. Admires(evan, paige) Basis

2. ~Married(beth,matt)	Assumption
3. Admires(evan, paige)	Reit: 1
4. (~Married(beth,matt)→Admires(evan,paige))	→-Intro:2-3

Recall that the rationale behind the →-Intro rule is that we may derive a conditional if we derive the consequent Q from the assumption of the antecedent P, and, perhaps, other sentences occurring earlier in the proof on wider margins. The defect of this rule, according to Anderson and Belnap is that "from" in "from the assumption of the antecedent P" is not taken seriously. They seem to have a point. By the lights of the →-Intro rule, we have derived line 4 but it is hard to see how we have derived the sentence at line 3 **from** the assumption at step 2 when we have simply reiterated the basis at line 3. Clearly, *~Married(beth,matt)* was not used in inferring *Admires(evan,beth)* at line 3. The relevance logician claims that the →-Intro rule in a correct natural deductive system should not make it possible to prove a conditional when the consequent was arrived at independently of the antecedent. A typical strategy is to use classes of numerals to mark the relevance conditions of basis sentences and assumptions and formulate the intro and elim rules to tell us how an application of the rule transfers the numerical subscript(s) from the sentences used to the sentence derived with the help of the rule. Label the basis sentences, if any, with distinct numerical subscripts. Let *a*, *b*, *c*, *etc*…range over classes of numerals. The → rules for a relevance natural deductive system may be represented as follows.

→-Elim

k. $(P \rightarrow Q)_a$
l. P_b
m. $Q_{a \cup b}$ →-Elim: k, l

→-Intro

k. $P_{\{k\}}$ Assumption
l. Q_b
m. $(P \rightarrow Q)_{b-\{k\}}$ →-Intro: k-l, provided $k \in b$

The numerical subscript of the assumption at line k must be new to the proof. This is insured by using the line number for the subscript.

In the directions for the →-Intro rule, the proviso that $k \in b$ insures that the antecedent P is used in deriving the consequent Q. Anderson and Belnap require that if the line *m* that results from the application of either rule is the conclusion of the proof the relevance markers be discharged. Here is a sample proof of the above two rules in action.

1. Admires(evan, paige)$_{1}$ Assumption
2. (Admires(evan, paige)→~Married(beth,matt))$_{2}$ Assumption
3. ~Married(beth,matt)$_{1,2}$ →-Elim: 1,2
4. ((Admires(evan,paige)→~Married(beth,matt))→~Married(beth,matt))$_{1}$ →-Intro: 2-3
5. (Admires(evan,paige)→((Admires(evan,paige)→~Married(beth,matt))→~Married(beth,matt))) →-Intro: 1-4

For further discussion see Anderson and Belnap (1962). For a comprehensive discussion of relevance deductive systems see their (1975). For a more up-to-date review of the relevance logic literature see Dunn (1986). We now consider the correctness of the ~-Elim rule and consider the rule in the rule in the context of using it along with the ~-Intro rule.

Here is a typical use in classical logic of the ~- Intro and ~-Elim rules. Suppose that we derive a contradiction from the assumption that a sentence P is true. So, if P were true, then a contradiction would be true which is impossible. So P cannot be true and we may infer that not-P. Similarly, suppose that we derive a contradiction from the assumption that not-P. Since a contradiction cannot be true, not-P is not true. Then we may infer that P is true by ~-Elim.

The intuitionist logician rejects the inference highlighted in italics. If a contradiction is derived from not-P we may infer that not-P is not true, i.e. that not-not-P is true, but it is incorrect to infer that P is true. Why? Because the intuitionist rejects the presupposition behind the ~-Elim rule, which is that for any proposition P there are two alternatives: P and not-P. The grounds for this are the intuitionistic conceptions of truth and meaning.

According to intuitionistic logic, truth is an epistemic notion: the truth of a sentence P consists of our ability to verify it. To assert P is to have a proof of P, and to assert not-P is to have a refutation of P. This leads to an epistemic conception of the meaning of logical constants. The meaning of a logical constant is characterized in terms of its contribution to the criteria of proof for the sentences in which it occurs. Compare with classical logic: the meaning of a logical constant is semantically characterized in terms of its contribution to the determination of the truth conditions of the sentences in which it occurs. For example, the classical logician accepts a sentence of the form P v Q only when she accepts that at least one of the disjuncts is true. On the other hand, the intuitionistic logician accepts P v Q only when she has a method for proving P or a method for proving Q. But then the Law of Excluded Middle no longer holds, because a sentence of the form *P or not-P* is true, i.e. assertible, only when we are in a position to prove or refute P, and we lack the means for verifying or refuting all sentences. The alleged problem with the ~ -Elim rule is that it illegitimately extends the grounds for asserting P on the basis of not-not-P since a refutation of not-P is not *ipso facto* a proof of P.

Since there are finitely many McKeons and the predicates of M seem well defined, we can work through the domain of the McKeons to verify or refute any M-sentence and so there doesn't seem to be an M-sentence that is neither verifiable not refutable. However, consider a language about the natural numbers. The open sentence 'x=y +z' is decidable for all x, y, and z. This is to say that for any natural numbers x, y, and z, we have an effective procedure for determining whether or not x is the sum of y and z. Hence, we may assert that for all x, y, and z, either x =y+z or not. Let 'A(x)' abbreviate 'if x is even and greater than 2 then there exists primes y and z such that x=y+z'. Since there are algorithms for determining of any number whether or not it is even, greater than 2, or prime, that the predicate 'A' holds for a given natural number x is decidable for we can effectively determine for all pairs of numbers less than x whether they are prime. However, there is no known method for verifying or refuting Goldbach's conjecture, *for all x, A(x)*. Even though, for each *x*, 'A(x) or not-A(x)' is decidable, 'for all x, A(x)' is not. That is, we are not in a position to hold that either Goldbach's conjecture is true or it is not. Clearly, verification of the conjecture *via* an exhaustive search of the domain of natural numbers is not possible since the domain is non-finite. Minus a counterexample or proof of Goldbach's conjecture, the intuitionist demurs from asserting that either Goldbach's conjecture is true or it is not. This is just one of many examples where the intuitionist thinks that the law of excluded middle fails.

In sum, the legitimacy of the ~-Elim rule requires a realist conception of truth as verification transcendent. On this conception, sentences have truth-values independently of the possibility of a method for verifying them. Intuitionistic logic abandons this conception of truth in favor of an epistemic conception according to which the truth of a sentence turns on our ability to verify it. Hence, the inference rules of an intuitionistic natural deductive system must be coded in such a way to reflect this notion of truth. For example, consider an intuitionistic language in which *a*, *b*,… range over proofs, '*a*: P' stand for '*a* is a proof of P', and '(*a*, *b*)' stand for some suitable pairing of the proofs *a* and *b*. The &-rules of an intuitionistic natural deductive system may look like the following.

&-Intro

k. *a*: P
l. *b*: Q
m. (*a*,*b*): (P & Q) &-Intro: k, l.

& -Elim

k. (*a*,*b*): (P & Q)
l. *a*: P & Elim: k

k. (*a*,*b*): (P & Q)
l. *b*: Q & Elim: k

Apart from the negation rules, it is fairly straightforward to dress the intro and elim rules of N with a proof interpretation as is illustrated above with the &- rules. For the details see Van Dalen (1999). For further introductory discussion of the philosophical theses underlying intuitionistic logic see Read (1995) and Shapiro (2000). Tennant (1997) offers a more comprehensive discussion and defense of the philosophy of language underlying intuitionistic logic.

We now turn to the ∃-Intro and ∀-elim rules. Consider the following two inferences.

(1) Evan is male
∴ (2) Some are male

(3) All are male
∴ (4) Evan is male

Both are correct by the lights of our system N. Specifically, the M-sentence that corresponds with (2) is derivable from the M-sentence correlate of (1) by the ∃-Intro rule and we get the M correlate of (4) from the M correlate of (3) by the ∀-elim rule. Note an implicit assumption required for the legitimacy of these inferences: every individual constant refers to an element of the quantifier domain. If this existence assumption, which is built into the semantics for M and reflected in the two quantifier rules, is rejected, then the inferences are unacceptable. What motivates rejecting the existence assumption and denying the correctness of the above inferences?

Well, there are contexts in which singular terms are used without assuming that they refer to existing objects. For example, it is perfectly reasonable to regard the individual constants of a language used to talk about myths and fairy tales as not denoting existing objects. For example, it seems inappropriate to infer that some actually existing individual is jolly on the basis that the sentence *Santa Claus is jolly* is true. Also, the logic of a language used to debate the existence of God should not presuppose that *God* refers to something in the world. The atheist doesn't seem to be contradicting herself in asserting that God does not exist. Furthermore, there are contexts in science where introducing an individual constant for an allegedly existing object such as a planet or particle should not require the scientist to know that the purported object to which the term allegedly refers actually exists. A logic that allows non-denoting individual constants (terms that do not refer to existing things) while maintaining the existential import of the quantifiers ('∀x' and '∃x' mean something like 'for all existing individuals x' and 'for some existing individuals x', respectively) is called a free logic. In order for the above two inferences to be correct by the lights of free logic, the sentence *Evan exists* must be added to the basis. Correspondingly, the ∃-Intro and ∀-elim rules in a natural deductive system for free logic may be portrayed as follows. Again, let 'Ωv' be a formula in which 'v' is the only free variable, and let 'n' be any name.

∀-Elim		∃-Intro	
k. ∀vΩv		k. Ωn	
l. E!n		l. E!n	
m Ωn	∀ Elim: k, l	m. ∃vΩv	∃-Intro: k, l

E!n abbreviates *n exists* and so we suppose that *E!* is an item of the relevant language. The ∀-Intro and ∃-Elim rules in a free logic deductive system also make explicit the required existential presuppositions with respect to individual constants (for details see Bencivenga (1986), p. 387). Free logic seems to be a useful tool for representing and evaluating reasoning in contexts such as the above. Different types of free logic arise depending on whether we treat terms that do not denote existing individuals as denoting objects that do not actually exists or as simply not denoting at all.

In sum, there are contexts in which it is appropriate to use languages whose vocabulary and syntactic, formation rules are independent of our knowledge of the actual existence of the entities the language is about. In such languages, the quantifier rules of deductive system N sanction incorrect inferences, and so at best N represents correct deductive reasoning in languages for which the existential presupposition with respect to singular terms makes sense. The proponent of system N may argue that only those expressions guaranteed a referent (e.g., demonstratives) are truly singular terms. On this view, advocated by Bertrand Russell at one time, expressions that may not have a referent such as S*anta Claus*, *God*, *Evan, Bill Clinton, the child abused by Michael Jackson,* are not genuinely singular expressions. For example, in the sentence *Evan is male*, *Evan* abbreviates a unique description such as *the son of Matt and Beth*. Then *Evan is male* comes to

There exists a unique x such that x is a son of Matt and Beth and x is male.

From this we may correctly infer that some are male. The representation of this inference in N appeals to both the ∃-Intro and ∃-Elim rules, as well as the &-Elim rule. However, treating most singular expressions as disguised definite descriptions at worst generates counter-intuitive truth-value assignments (*Santa Claus is jolly* turns out false since there is no Santa Claus) and seems at best an unnatural response to the criticism posed from the vantage point of free logic.

For a short discussion of the motives behind free logic and a review of the family of free logics see Read (1995), Chapter 5. For a more comprehensive discussion and a survey of the relevant literature see Bencivenga (1986). Morscher and Hieke (2001) is a collection of recent essays devoted to taking stock of the past fifty years of research in free logic and outlining new directions.

This completes our discussion of the status of the characterization of the logical consequence relation in terms of deducibility in system N. In sum, the defense of inferentialism against Prior's criticism has uncovered that N's rules for '=' and '~' fail to exhibit harmony, and, therefore, the elim-rules cannot be justified in terms of the meanings of the corresponding expressions. This constitutes a criticism of N, since it is plausible to think that an expression's inference rules are derived from its meaning. Tarski argues that since logical consequence is not compact according to the common concept, it cannot be defined in terms of deducibility in a deductive system. So, there are intuitively correct principles of inference that cannot be represented in deductive systems such as N (e.g., the ω-rule). Is N correct as far as it goes? In other words: Do the introduction and elimination rules of N represent correct principles of inference? We sketched three motives for answering in the negative, each leading to a logic that differs from the classical one developed here and which requires altering intro and elim rules of N.

Bibliography

Almog, J. (1989): "Logic and the World" pp. 43-65 in *Themes From Kaplan*, ed. Almog, J., Perry, J., and Wettstein, H., New York: Oxford UP (1989).

Alward, P. (2005): "Between the Lines of Age: Reflections on the Metaphysics of Words", *Pacific Philosophical Quarterly* 86, pp. 172-187.

Anderson, A.R., Belnap, N. (1962): "Entailment" pp. 76-110 in *Logic and Philosophy*, ed. by Iseminger, G., New York: Appleton-Century-Crofts, (1968).

Anderson, A.R., Belnap, N. (1975): *Entailment: The Logic of Relevance and Necessity*, Princeton: Princeton University Press.

Aristotle: *Basic Works*, ed. By R. McKeon, New York: Random House, (1941).

Bach, K. (2000): "Quantification, Qualification, and Context A Reply to Stanley and Szabo", *Mind & Language*, 15, pp. 262-283.

Barwise, J. and Etchemendy J. (2002): *Language, Proof and Logic*. Chicago: University of Chicago Press and CSLI Publications.

Beall, J.C. (2004): "Introduction: At the Intersection of Truth and Falsity" pp. 1-19 in *The Law of Non-Contradiction: New Philosophical Essays*, Oxford: Clarendon Press, (2004).

Belnap, N. (1962): "Tonk, Plonk, and Plink", *Analysis* 22, pp. 130-134.

Bencivenga, E. (1986): "Free logics" pp. 373-426 in Gabbay and Geunthner (1986).

——. (1999): "What is Logic About?" pp. 5-19 in Varzi (1999).

Boolos, G. (1993): "Whence the Contradiction?", Aristotelian Society Supplementary Volume 67, pp. 213-233.

Cartwright, R. (1994): "Speaking of Everything", *Nous* 28, pp. 1-20.

Clark, P. (1993): "Sets and Indefinitely Extensible Concepts and Classes", *Aristotelian Society Supplementary Volume* 67, pp. 235-249.

Dummett, M. (1963): "The Philosophical Significance of Gödel's Theorem", *Ratio* 5, pp. 140-155. Also, pp.186-201 in *Truth and Other Enigmas*. Cambridge, MA: Harvard UP (1978). Page references are to the later.

——. (1981): *Frege: Philosophy of Language*, second ed. Cambridge, MA: Harvard UP.

——.(1991a): *The Logical Basis of Metaphysics*. Cambridge, MA: Harvard UP.

——. (1991b): "Does Quantification Involve Identity?" in *Peter Geach: Philosophical Encounters*. Deventer. Also, pp. 308-327 in *The Seas of Language*. Oxford: Clarendon Press (1993). Page references are to the later.

——. (1991b): *Frege: Philosophy of Mathematics*. Cambridge, MA: Harvard UP.

Dunn, M. (1986): "Relevance Logic and Entailment" pp. 117-224 in Gabbay and Geunthner (1986).

Epstein, R. (2001): *Predicate Logic: The Semantics Foundations of Logic*. Belmon, CA: Wadsworth.

Etchemendy, J. (1983): "The Doctrine of Logic as Form", *Linguistics and Philosophy* 6, pp. 319-334.

——. (1988): "Tarski on truth and logical consequence", *Journal of Symbolic Logic* 53, pp. 51-79.

——. (1999): *The Concept of Logical Consequence.* Stanford: CSLI Publications.

——. (2008): "Reflections on Consequence" pp. 263-299 in *New Essays on Tarski and Philosophy* , ed. by D. Patterson. Oxford: Oxford UP (2008).

Fine, K. (1994): "Essence and Modality" pp. 1-16 in *Philosophical Perspectives VIII: Logic and Language*, ed. James Tomberlin. Atascadero, CA: Ridgeview, (1994).

Fitch, F.B. (1952): *Symbolic Logic: An Introduction*. New York: The Ronald Press Co.

Forbes, G. (1994): *Modern Logic: A text in Elementary Symbolic Logic*. Oxford: Oxford UP.

Fraassen van, B. (1969): "Presuppositions Supervaluations and Free Logic", pp. 67-91 in *The Logical Way of Doing Things*, ed. by K. Lambert. New Haven, CT: Yale UP (1969).
Frege, G. (1950): The Foundations of Arithmetic. Trans. By J. L. Austin. Evanston: Northwestern UP.
Gabbay, D. and F. Guenthner, eds (1983): *Handbook of Philosophical Logic*. Vol 1, Dordrecht: D. Reidel Publishing Company.
–––. (1986): *Handbook of Philosophical Logic.*Vol. 3, Dordrecht: D. Reidel Publishing Company.
Geach, P.T. (1962): *Reference and Generality*. Cornell: Cornell UP.
Gentzen, G. (1935): "Investigations Into Logical Deduction", pp. 68-128 in *Collected Papers*,ed. by M.E. Szabo. Amsterdam: North-Holland Pub. Co., (1969).
Goldfarb, W. (2003): *Deductive Logic*. Indianapolis, IN: Hackett, Pub. Co.
Gorman, M. (2005): "The Essential and the Accidental", *Ratio* 18, pp. 276-289.
Haack, S. (1978): *Philosophy of Logics* , Cambridge: Cambridge University Press.
–––. (1996): *Deviant Logic, Fuzzy Logic*, Chicago: The University of Chicago Press.
Hamlyn, D.W. (1967): "A Priori and A Posteriori" pp.105-109 in *The Encyclopedia of Philosophy* Vol. 1 ed. by P. Edwards, New York: Macmillan Pub. Co., Inc. & The Free Press, (1967).
Hanson, W. (1997): "The Concept of Logical Consequence", *The Philosophical Review* 106, pp. 365-409.
Henkin, L. (1967): "Formal Systems and Models of Formal Systems", pp. 61-74 in *The Encyclopedia of Philosophy* Vol. 8, ed. by P. Edwards. New York: Macmillan Pub. Co., Inc. & The Free Press, (1972).
Hjortland Thomassen, O. (2007): "Proof-theoretic harmony and structural assumptions", unpublished. Available at http://olethhjortland.googlepages.com/Cambridge.pdf.
Hodges, W. (1983): "Elementary Predicate Logic" in Gabbay, D. and F. Guenthner (1983).
–––. (1986): "Truth in a Structure", *Proceedings of Aristotelian Society* 86, pp. 135-151.
–––. (2007): "The scope and limits of logic" pp. 41-63 in *Handbook of the Philosophy of Science: Philosophy of Logic*, ed. by D. Jacquette. Amsterdam: Elsevier, (2007).
Hughes, G. E. and M.J Cresswell (1996): *A New Introduction to Modal Logic*, London: Routledge.
Hunter, G. (1971): *Metalogic: An Introduction to the Metatheory of Standard First Order Logic*, Berkeley: University of California Press.
Jané, I. (2006): "What is Tarski's Common Concept of Logical Consequence?" *The Bulletin of Symbolic Logic* 12, pp. 1-42.
Kneale, W. (1961): "Universality and Necessity", *British Journal for the Philosophy of Science* 12, pp. 89-102.
Koslicki, K. (1997): "Isolation and Non-Arbitrary Division: Frege's Two Criteria for Counting", *Synthese* 112, pp. 403-430.
Koslow, A. (1999): "The Implicational Nature of Logic: A Structuralist Account" pp. 111-155 in Varzi (1999).
Lavine, S. (2006): "Something About Everything" pp. 98-148 in Rayo and Uzquiano (1996).
Lowe, E.J. (1997): "Objects and Criteria of Identity" pp. 613-633 in *A Companion to the Philosophy of Language*, ed. by B. Hale and C. Wright. Oxford: Blackwell (1997).
–––. (2003): "Individuation" pp. 75-95 in The Oxford handbook of Metaphysics, ed. by M. Loux and D. Zimmerman. Oxford; Oxford UP (2003).
Martin, J.N. (1987): *Elements of Formal Semantics: An Introduction to Logic for Students of Language*. Orlando: Academic Press, Inc.
Martin, R.M. (1962): "Existential Quantification and the 'Regimentation' of Ordinary Language", *Mind* 71, pp.525-529.

McCarthy, T. (1981): "The Idea of a Logical Constant" *Journal of Philosophy* 78, pp. 499-523.

——. (1998): "Logical constants" pp. 599-603 in *Routledge Encyclopedia of Philosophy* vol. 5, E. Craig ed., London: Routledge, (1998).

McGee, V. (1999): "Two Problems with Tarski's Theory of Consequence", *Proceedings of the Aristotelean Society* 92, pp. 273-292.

Milne, P. (1994): "Classical Harmony: Rules of Inference and the Meaning of the Logical Constants" *Synthese* 100, pp. 49-94.

Morscher E. and A. Hieke, eds. (2001): *New Essays in Free Logic: In Honour of Karel Lambert*. Dordrecht: Kluwer Academic Publishers.

Peacocke, C. (1976): "What is a Logical Constant?" *Journal of Philosophy* 73: 221-240.

Pelletier, F.J. (1999): "A History of Natural Deduction and Elementary Logic Textbooks" pp.105-138 in *Logical Consequence: Rival Approaches.* Ed. by J. Woods and B. Brown. Oxford: Hermes Science Publishing, (2001).

Peters, S. and D. Westerståhl (2006): *Quantifiers in Language and Logic*. Oxford: Clarendon Press.

Potter, M. (2004): *Set Theory and its Philosophy*. Oxford: Oxford UP.

Pospesel, H. (1976): *Introduction to Predicate Logic*. Englewood Cliffs, NJ: Prentice Hall.

Prawitz, D.(1965): *Natural Deduction: A Proof-Theoretical Study*. Stockholm: Almqvist & Wiksell.

——. (1977): "Meaning and proofs: on the conflict between classical and intuitionistic logic" *Theoria* 43, pp. 1-40.

Priest, G. (2001): An Introduction to Non-Classical Logic. Cambridge: Cambridge UP.

——. (2002): Beyond the Limits of Thought. Oxford: Clarendon Press.

Prior, A. (1960): "The Runabout Inference –Ticket" *Analysis* 21, pp. 38-39.

——. (1976): "What is Logic?" pp. 122-129 in *Papers in Logic and Ethics*, ed. by P.T. Geach and A.Kenny. Amherst: University of Massachusetts Press, (1976).

Putnam, H. (1969): "Is Logic Empirical?" in *Boston Studies in the Philosophy of Science* Vol. 5, ed. by M. Cohen and M.R. Wartofsky. New York: Humanities Press, (1969).

——. (1971): *Philosophy of Logic.* New York: Harper & Row.

Quine, W.V. (1960): *Word and Object*. Cambridge, MA: MIT Press.

——. (1986): *Philosophy of Logic*. 2nd ed., Cambridge: Harvard UP.

Ray, G. (1996): "Logical Consequence: A Defense of Tarski", *Journal of Philosophical Logic* 25, pp. 617-677.

Rayo, A, and G. Uzquiano, editors (2006): *Absolute Generality.* Oxford: Clarendon Press.

Read, S. (1995): *Thinking About Logic.* Oxford: Oxford UP.

——. (2000): "Harmony and Autonomy in Classical Logic" *Journal of Philosophical Logic* 29, pp. 123-154.

——. (2004): "Identity and Harmony" *Analysis* 64, pp. 113-119.

——. (2008): "Harmony and Modality" pp. 285-303 in *Dialogues, Logics and Other Strange Things: Essays in Honour of Shahid Rahman*, ed. by C. Dégremont, L. Kieff and H. Rückert, London: King's College Publications, (2008).

Restall, G. (2002): "Carnap's Tolerance, Meaning, And Logical Pluralism" *Journal of Philosophy* , pp. 426-443.

Russell, B. (1906): "On some difficulties in the Theory of Transfinite Numbers and Order Types", *Proceedings of the London Mathematical Society,* series 2, vol. 4. Also, pp. 29-53 in *Essays in Analysis*, ed. D. Lackey. London, (1973) pp. 135-64. Page references are to the later.

——. (1919): *Introduction to Mathematical Philosophy*, London: Routledge, 1993 printing.

Sainsbury, R.M. (1995): *Paradoxes,* second edition. Cambridge: Cambridge UP.

Shapiro, S. (1991): *Foundations without Foundationalism: A Case For Second-order Logic*, Oxford: Clarendon Press.

——. (1993): "Modality and Ontology" *Mind* 102, pp. 455-481.

——. (1998): "Logical Consequence: "Models and Modality" pp. 131-156 in *The Philosophy of Mathematics Today*, ed. by Matthias Shirn. Oxford, Clarendon Press (1998).

——. (2000): *Thinking About Mathematics* , Oxford: Oxford University Press.

——. (2005a): "Logical Consequence, Proof and Model-Theories" pp. 651-670

——. (2005b) *The Oxford Handbook of Philosophy of Mathematics and Logic*.

Shapiro S. and C. Wright (2006): "All Things Indefinitely Extensible" pp. 255-304 in Rayo and Uzquiano (2006).

Sher, G. (1989): "A Conception of Tarskian Logic" *Pacific Philosophical Quarterly* 70, pp.341- 368.

——. (1991): *The Bounds of Logic: A Generalized Viewpoint*, Cambridge, MA: The MIT Press.

——. (1996): "Did Tarski commit 'Tarski's fallacy'?" *Journal of Symbolic Logic* 61, pp.653-686.

——. (1999): "Is Logic a Theory of the Obvious?" pp.207-238 in Varzi (1999).

Smiley, T. (1995): "A Tale of Two Tortoises" *Mind* 104, pp. 725-36.

——. (1998): "Consequence, Conceptions of" pp. 599-603 in *Routledge Encyclopedia of Philosophy* vol. 2, ed. E. Craig. London: Routledge, (1998).

Stevenson, J.T. (1961): "Roundabout the Runabout Inference-Ticket" *Analysis* 21, pp. 124-128.

Sundholm, G. (1983): "Systems of Deduction" in Gabbay, D. and F. Guenthner.

Tarski, A. (1933): "*Pojecie prawdy w jezykach nauk dedukeycyjnych*" translated as "On the Concept of Truth in Formalized Languages" pp. 152-278 in Tarski (1983).

——. (1936): "On the Concept of Logical Consequence" pp. 409-420 in (1983). Also, Translated as "On the Concept of Following Logically". Trans. by M. Stroińska and D. Hitchcock (2002) *History and Philosophy of Logic* 23, pp. 155-196.

——. (1941): *Introduction to Logic and the Methodology of Deductive Sciences*. (1995) printing. New York: Dover Publications.

——. (1944): "The Semantic Conception of Truth and the Foundations of Semantics" *Philosophy and Phenomenological Research* 4, pp. 341-375.

——. (1983): *Logic, Semantics, Metamathematics* 2nd ed. Indianapolis: Hackett Publishing.

——. (1986): "What are logical notions?" *History and Philosophy of Logic* 7, pp. 143-154.

Tharp, L. (1975): "Which Logic is the Right Logic?" *Synthese* 31, pp. 1-21.

Tennant, N. (1978): *Natural Logic*, Edinburgh: Edinburgh UP.

——. (1997): *The Taming of the True*, Oxford: Clarendon Press.

——. (2007): "Inferentialism, Logicism, Harmony, and a Counterpoint" in *Essays for Crispin Wright* ed. by A. Miller and A. Coliva. Oxford: Oxford UP

Warbrod, K., (1999): "Logical Constants" *Mind* 108, pp.503-538.

Williamson, T. (1999): "Existence and Contingency", *Proceedings of the Aristotelian Society* Supplementary Vol. 73, 181-203.

Wright, C. (1987): *Frege's Conception of Numbers as Objects.* Aberdeen: Aberdeen UP.

——. (1999): "Is Hume's Principle Analytic", *Notre Dame Journal of Formal Logic* 40, pp.6-30. Also pp. 307-332 in B. Hale and C. Wright *The Reason's Proper Study* Oxford: Oxford UP, (2001). Page references are to the later.

Van Dalen, D. (1999): "The Intuitionistic Conception of Logic", pp. 45-73 in Varzi (1999).

Varzi, A., ed. (1999): *European Review of Philosophy, Vol. 4: The Nature of Logic*, Stanford: CSLI Publications.