This article appears in <u>The Blackwell Guide to Philosophical Logic</u>, *Lou Goble (ed), Blackwell Publishers 2001, pp 115-135.*

<u>Logical Consequence</u>

<u>Patricia A. Blanchette</u>

<u>Introduction</u>

Whenever one asserts a claim of any kind, one engages in a commitment not just to that claim itself, but to a variety of other claims that follow in its wake, claims that, as we tend to say, *follow logically* from the original claim. To say that Smith and Jones are both great basketball players is to say something from which it follows that Smith is a great basketball player, that someone is a great basketball player, that there is something at which Smith is great, and so on.

This general fact, that certain claims follow logically from others, is the central concern of a theory of logical consequence. Logical consequence is just the relation that connects a given claim or set of claims with those things that follow logically from it: to say that B is a logical consequence of A is simply to say that B follows logically from A. All of our ordinary reasoning turns on the recognition of this relation. When we notice, for example, that a certain prediction follows from a given theory, that a particular view is a consequence of some initial commitments, that a collection of premises entails a given conclusion, and so on, we are engaged in reasoning about logical consequence.

The other logical properties and relations whose recognition is central to ordinary reasoning are closely related to, and can be defined in terms of, logical consequence. We say that an argument is *valid* iff its conclusion is a logical consequence of its premises; a set of claims □ *entails* a claim □ iff □ is a logical consequence of □; a set of claims is *consistent* iff no contradiction is a logical consequence of it, and a claim □ is *independent* of a set of claims □ iff □ is not a logical consequence of □. Finally, a claim is a *logical truth* iff it is a logical consequence of the empty set of claims.

The investigation of logical consequence and related notions consists largely in the attempt (a) to give a systematic treatment of the *extension* of this relation, i.e. of the issue of *which* claims do in fact follow logically from which others; and (b) to give an informative account of the *nature* of the relation. There is much room for debate about both of these issues. Though there is no doubt about the fact that some claims do in fact follow logically from others, and (perhaps more obviously) that some sets of beliefs are inconsistent, that some arguments are definitely not valid, and so on, there is room for disagreement about cases. Philosophers have disagreed, for example, about whether certain purely mathematical claims are logical consequences of apparently non-mathematical claims. They have disagreed about whether existential claims about properties follow logically from ordinary predications. And so on.

To give a determinate answer to each and every question of the form "Does this follow logically from that?" will require, among other things, a decision about the precise boundaries separating logical consequence from set-theoretic or mathematical fact, and a decision about the metaphysical commitments of various kinds of claims. The project of clarifying the extension of the logical-consequence relation thus turns to some extent on issues outside the scope of the philosophy of logic proper, on issues, for example, that fall within the camp of pure metaphysics, philosophy of mathematics, and related fields. But some of the central questions about the extent of the relation depend for their answers on the second of the two topics noted above, namely on the issue of the *nature* of logical consequence. Reflection on some clear, easily-recognized cases of logical consequence (e.g. that *Socrates is mortal* follows logically from the pair of claims *All humans are mortal* and *Socrates is human*) easily reveals some straightforward necessary conditions for logical consequence. It is uncontroversial, for example, that the logical consequences of true claims must themselves be true. That is, as we can put it briefly, that the relation of logical consequence is *truth-preserving*. But things get more difficult, and more contentious, when we try to fill in more details. Some of the disagreements here turn, for example, on questions about whether the conclusion of a valid argument must in some sense be "about" the same subject-matter as are its premises, and about whether there is a clear sense in which the fundamental principles of logic must hold independently of any particular subject-matter, to be, as it is often put, "topic-neutral." Further disagreements arise about whether, and in what sense, logical truths must always be *necessary* truths, and so on. Different answers to these questions will deliver different views about the precise extension of the logical-consequence relation, and will consequently give rise to different systematic treatments of consequence in the form of formal systems of logic.

The purpose of this chapter is to provide a brief introduction to the central issues surrounding the nature and the extension of logical consequence, and to the role of formal systems in the investigation of consequence.

Early Formal Systems and Accuracy

A *formal system* of logic consists of a rigorously-specified *formal language* (i.e., a collection of formulas defined solely in terms of their syntax) together with a *deductive system* (i.e., a specification of those series of formulas that will count as *deductions* of particular formulas from collections of formulas). Since the end of the nineteenth century, formal systems have been widely used as means of codifying and analyzing the relation of logical consequence and its associated notions.

The earliest formal systems, e.g. those of Gottlob Frege (1848-1925) were intended to give a way of rigorously demonstrating relations of logical consequence – i.e., of demonstrating of particular claims that they were indeed logical consequences of particular sets of claims. The intention in designing the system was, to put it somewhat loosely, that the system would include a *deduction* of a formula □ from a set □ of formulas only if □ was indeed a logical consequence of □. Deducibility within the system was to have been a reliable indicator of logical consequence.

The "somewhat loose" character of the above description is due to the fact that that the formulas themselves – i.e. strings of marks on paper – are not in fact the items that bear the logical consequence relation to one another. Thus we cannot, strictly speaking, describe the goal of system-design as one of making sure the deducibility-relation is included in the logical-consequence relation. As Frege saw it, the items that bear logical relations to one another are nonlinguistic propositions, the kinds of things that are expressed by fully-interpreted sentences, and that are the objects of the propositional attitudes. And, as Frege saw it, his formulas as they appeared in deductions always expressed determinate propositions. Thus the Fregean goal for an adequate formal system can be described, now accurately, as follows: a formula □ is to be deducible from a set □ of formulas only if the proposition expressed by □ is a logical consequence of the propositions expressed by the members of □.[1]

It is, in principle at least, a straightforward matter to check whether a system satisfies this requirement. Since the *deducibility* relation is typically defined in the familiar way in terms of axioms and rules of inference, the quality-control check is (in principle, at least) simple: one checks to see that every proposition expressible by an axiom is a truth of logic, and that each rule of inference countenances the deduction of a formula □ from an n-tuple of formulas $\square_1 \ldots \square_n$ only if the proposition expressed by □ is a logical consequence of those expressed by $\square_1 \ldots \square_n$.

Frege himself did not offer any uniform method for carrying out this "checking", i.e. for demonstrating that a given proposition is in fact a truth of logic, or that a given proposition does in fact follow logically from a particular collection of propositions. He seems to presume that the very simplest cases of logical truth and of logically-valid inference are obvious when encountered. The accuracy of the formal system was to have been established by simply pointing out that the propositions expressed by axioms were indeed obvious truths of logic, and that the rules of inference obviously generated only logical consequences from premises. The importance of the formal system was, in large part, that once an audience had granted the handful of (ostensibly) immediately-obvious logical principles required for axioms and rules of inference, it was a straightforward matter to demonstrate the validity of considerably complicated and non-obvious arguments. The accuracy of the system as a whole rested simply on the logical status of the axioms and rules of inference, and the validity

[1] See Frege's *Grundgesetze* Vol I, esp the Introduction and §§14, 15, 18, 20.

of extremely complicated arguments was guaranteed by the accuracy of the system.

As it turned out, Frege's favored formal system (that of the *Grundgesetze der Arithmetik* of 1893/1903) was not in fact accurate. His deducibility relation contained a subtle but important flaw, with the result that the system contains deductions of both a formula □ and its negation ~□ from the empty set of premises. And since the propositions expressed by such pairs of formulas cannot both be logical truths, Frege's deducibility relation cannot be considered a reliable indicator of logical consequence. Some of the fundamental claims that Frege took to be "obvious" logical truths were in fact falsehoods.[2] Successors to Frege's system of course avoid this particular error, and a number of them offer presumably-accurate indications of logical consequence. Before turning to a discussion of contemporary formal systems, two features of Frege's approach to formal systems are worth noting.

The first is the Fregean view of the *bearers* of the logical relations. Of central concern in the philosophy of logic, the question is this: what kinds of things, exactly, are the logical truths, the relata of the logical-consequence relation, the components of valid arguments, and so on? Frege's answer is, as noted above, that these bearers of logical relations are a particular kind of abstract object, namely, nonlinguistic *propositions*. The attraction of this view stems from the fact that these propositions are also, as Frege and most proposition-theorists see it, both the semantic values of our utterances and the objects of our propositional attitudes. The combination of these views gives an easy explanation of the fact that not only our assertions, but also our beliefs, the contents of our hopes and fears, and so on, can have logical implications. The view also helps to make sense of the apparent logical connections between

[2] The problem arises from Frege's assumption that to every predicate or open sentence there corresponds a set-like entity called an *extension*. As Russell's Paradox shows, this assumption is false. See Russell's letter to Frege of 16 June 1902 and Frege's response of 22 June 1902 (both translated and printed in *Philosophical and Mathematical Correspondence* pp 130-133). See also Frege, *Grundgesetze* Vol II, Appendix (trans in *Basic Laws of Arithmetic* at pp 127-141).

these different kinds of entities: the very thing that forms the content of one person's desire can logically contradict the content of another's assertion and of yet a third person's belief; and the straightforward explanation of this, on the propositional view, is that the desire, the assertion, and the belief all have propositions as their contents. But the view is not without problems. The central difficulty with the view of nonlinguistic propositions as the bearers of the logical relations is that, arguably, it is doubtful that there are such things as nonlinguistic propositions. Reasons for doubting the existence of propositions stem primarily from the difficulty of giving clear criteria of individuation for propositions, from general worries about abstract objects, and from considerations of ontological parsimony.[3]

Skepticism about nonlinguistic propositions leads to the alternative view that *sentences* are the bearers of the logical relations. Some caution is required here, however, about precisely what is meant by the claim that sentences are logical truths, are logical consequences of one another, and so on. Taking a sentence to be simply a series of marks or sounds, the view that the logical relations are borne by sentences is untenable. The string of symbols "All men are mortal" does not, by itself, have any logical implications, any more than does the string "Axz%f." Though it is tempting to view the first string as having a rich collection of logical implications, it is important to note that this temptation is felt only when we take the sentence to be not merely a string of shapes on paper, but rather to be something with a determinate *meaning*. A bare series of symbols does not have any logical properties at all, though a string of symbols together with the right kind of semantic value certainly does. The view that the logical relations obtain between *sentences*, then, is only a reasonable view if by "sentence" we mean something like "series of symbols together with a determinate meaning." The view, in short, is that the bearers of the logical relations are meaningful sentences.

The two views in question (that nonlinguistic propositions are the primary bearers of the logical relations, and that sentences together with their meanings

[3] See: Quine "On What There Is," and *Philosophy of Logic* Ch. 1; also Cartwright, "Propositions."

are the primary bearers of the logical relations) are not importantly different if, as is sometimes done, one takes the meanings of sentences just to be nonlinguistic propositions. But if one takes the meaning of a sentence to be something non-propositional, for example a pattern of use in a given population, then the two views are importantly different, with the first but not the second committed to the existence of something like Fregean propositions. The latter understanding of "meaning" is of course required by those whose view is motivated by skepticism about the existence of nonlinguistic propositions.

Despite the intrinsic interest of this issue, the difference between the two views (propositional and sentential) of the bearers of the logical properties and relations will not be terribly important in what follows, and we will not pause here to adjudicate between them. The important point about both views is that they are fundamentally "semantic" in the sense that they construe the logical relations as obtaining either between meanings themselves (propositions), or between pairs of syntactic items and meanings. And this is as it should be: as noted above, the logical relations do not obtain between bare syntactic items, but only between items which make some determinate claim on the world; which are, in brief, meaningful. In what follows, we shall simply continue our practice of referring to the relata of the logical relations as *claims*, taking the word to be ambiguous between nonlinguistic propositions and meaningful sentences. Except where noted, everything said below will apply on either reading.

The second feature worth noting here about the conception of logic underlying Frege's and similar formal work concerns the distinction between the pretheoretic relation of logical consequence and the various relations of formal deducibility given by particular formal systems. The relation of logical consequence is *pretheoretic* in the sense that neither the relation itself, nor our recognition of the relation, depends upon the existence or the deliverances of formal systems. Similarly for the related notions of logical truth, consequence, consistency, and so on. When we infer *Katy is wise* from the pair of claims *All of John's children are wise* and *Katy is a child of John's*, we recognize a connection between these claims that would have held whether or not anyone had ever invented formal systems, and whether or not any of those systems had pronounced the inference valid. Similarly for the ordinary notions of inconsistency, validity, entailment, and so on that we recognize in everyday reasoning. These logical properties and relations link our assertions, beliefs, and theories one to another in ways that do not depend upon the results of work in formal logic. The dependence is, rather, the other way round: standard systems of formal logic will count the argument just noted (or a formalized version of it) as valid because those systems are designed to accurately reflect the pretheoretic logical properties and relations. It is only with respect to this sense of a system-independent notion of logical consequence that we can make sense of the idea of a formal system's being accurate or inaccurate, since the accuracy of the system is a matter of the extent to which its relation of formal deducibility reliably indicates the pretheoretic relation of logical consequence. And, of course, it is only against the background of such a system-independent notion of logical consequence that we can agree with Frege's later assessment of his own formal system as inaccurate in the way noted above.

Contemporary Formal Systems

Contemporary formal systems differ from Frege's in two ways that are relevant to the issue of their accurate reflection of the logical properties and relations. First of all, we do not, these days, typically view the formulas that occur in deductions as each expressing unique claims. Each formula is typically thought, rather, to be capable of expressing a broad range of claims. The guidelines governing precisely *which* claims each formula can appropriately express are seldom made explicit; they are simply the rules of thumb we pass on when teaching students how to do "translations" between formal and natural languages. They are, in the typical case, rules about the fixed meanings to be assigned to the logical constants, the kinds of meanings assignable to the members of each syntactic category, rules of compositionality, and so on. These are the rules we have in mind when we note that, for example, "□xFx" can be used to formalize the claims *There's at least one prime* or *Someone is French*, but not *All cows are mammals*.

The second relevant difference between Fregean and standard contemporary systems is that the latter typically incorporate a *model-theoretic* apparatus. A *model* for a formal language is a function which, while meeting a variety of requirements specific to that language, assigns a truth-value to each closed formula of the language. The standard requirements include, for example, the requirement that a model assign *true* to a formula of the form (□ & □) only if it assigns *true* to both □ and □. For a quantified language, the assignment of truth-values proceeds via an assignment of individuals and sets to the atomic parts of formulas.

Instead of assessing the adequacy of formal systems in the Fregean way, by directly examining the relationship between deducible formulas and the claims they express, typical practice with contemporary systems is to assess the adequacy of a system by examining the relationship between deducible formulas and the truth-values assigned those formulas by various models. Where □ is a set of formulas of a formal system S and □ is a formula of S, we say that □ is a *model-theoretic consequence* in S of □ if every one of S's models that assigns *true* to each member of □ also assigns *true* to □. We abbreviate this claim as: "□ $|=_S$ □." We also say that a formula □ is a *model-theoretic* truth of S if every one of S's models assigns *true* to □. A central question that arises for a formal system with a model-theoretic apparatus is that of the coincidence between the relation of model-theoretic consequence and the relation of deducibility. Abbreviating "□ is deducible in S from □" as: "□ $|-_S$ □", the two halves of the coincidence-claim form the *soundness* and *completeness* theorems for S, as follows:

Soundness of S: For every set □ of formulas of S, and every formula □ of S: If □ $|-_S$ □, then □ $|=_S$ □.

Completeness of S: For every set □ of formulas of S, and every formula □ of S: If □ $|=_S$ □, then □ $|-_S$ □.

If one's primary interest is in devising a formal system whose deductive and model-theoretic consequence-relations coincide, then the soundness and completeness theorems are of interest in their own right. When, on the other hand, the purpose is the design of a formal system that will be a reliable indicator of logical consequence, these theorems are of interest largely because they allow us to infer the adequacy of one of these consequence-like relations (deducibility or model-theoretic consequence) from the other. If we know that model-theoretic consequence within a system S is a reliable indicator of logical consequence, then the soundness theorem for S will give us a reliability result for S's deducibility relation. If on the other hand we have an independent guarantee of the reliability of S's deducibility relation with respect to logical consequence, then the completeness theorem for S will establish the reliability in this regard of S's model-theoretic apparatus.

Because unlike their Fregean antecedents, standard contemporary systems take each formula to be capable of formalizing a wide range of claims, we need to introduce somewhat more complexity in formulating the questions of the reliability of the model-theoretic and the deductive consequence relations of formal systems. We cannot simply ask whether □'s being a model-theoretic consequence in S of □ entails that *the claim* expressed by □ is a logical consequence of *the set of claims* expressed by □, since there are no unique claim and set of claims expressed by □ and by □, respectively. We want, rather, to ask whether this implication holds for *all* of the claims and sets of claims expressible by □ and □ respectively. Similarly for the relation □ |-$_S$ □.

Consider, for example, the set of formulas {□x(Fx □ Gx), Fa} and the formula Ga. We can take these formulas to express, respectively, the claims *All primes are odd; Seven is prime;* and *Seven is odd*. On another occasion, we might use these formulas to represent the trio *All sheep are mammals; Dolly is a sheep;* and *Dolly is a mammal*. Each such assignment of claims to the formulas of a formal language is what we shall call a *reading* of that language. That is to say: a *reading* of a language is an assignment of claims to its closed formulas in a way that satisfies the usual "rules of thumb," of the kind mentioned above, for that language. When we ask whether deducibility in a given formal system S is a reliable indicator of logical consequence, we are interested in whether this reliability holds for *each* of the readings of the language. Similarly for the question of the reliability of the model-theoretic consequence relation for S.

Where □ and □ are a set of formulas and a formula, respectively, of the language of a formal system S, and R is a reading of that language, by "R(□)" we shall mean the set of claims assigned by R to □, and by "R(□)" the claim assigned by R to □. For example: where R1 is the first reading given in the above example, R1({□x(Fx □ Gx), Fa}) is the set of claims {*All primes are odd, Seven is prime*}, and R1(Ga) is *Seven is odd*. Our questions about the reliability of the deductive and model-theoretic consequence relations of a formal system S, then, can be expressed as the questions of whether it is generally true that:

(i) If □ $|-_S$ □, then R(□) is a logical consequence of R(□); and

(ii) If □ $|=_S$ □, then R(□) is a logical consequence of R(□).

If (i) holds for all □, □, and R for a formal system S, then we will say that S's deducibility relation is reliable. If (ii) holds for all □, □, and R for a system S, then we will say that S's model-theoretic consequence relation is reliable.

It is sometimes assumed that, at least for those formal systems standardly in use, the model-theoretic consequence relation is "automatically" a reliable indicator of logical consequence, which is to say that (ii) is obviously satisfied by such systems. This assumption tends to rest on the view that the relation of model-theoretic consequence is merely a tidied-up version of, or a successful re-description of, the pretheoretic relation of logical consequence itself. If this is the case, then the question of the reliability of the deducibility relation is immediately reducible to the question of its satisfaction of the soundness theorem. This assumption of the coincidence between model-theoretic consequence and logical consequence has, however, been challenged, and the grounds for inferring deductive reliability from soundness are by no means obvious.[4]

Satisfaction of (i) and (ii) are two of the central issues to be treated in establishing the accuracy of a formal system that has both a deductive and a model-theoretic apparatus. If we are interested in the use of the system not only to give positive judgments of logical consequence, but also negative such judgments, we will be interested in the stronger biconditional versions of (i) and (ii).

The central difficulty in establishing either (i) or (ii) (or their strengthened biconditional versions) is that we have no independent test for satisfaction of the consequent of each. After all, if we already had a reliable test of logical consequence, we would not be in the position of devising formal systems to provide such a test. Nevertheless, we can in fact formulate some relatively straightforward necessary conditions on logical consequence, some of which can be used to give at least partial evaluations of formal systems. The idea, roughly, is that if C is some condition that must be met by the logical-consequence relation, then no formal system whose deductive/model-theoretic consequence-relation fails to meet C can be said to satisfy (i)/(ii). We take a look below at a small sample of such conditions.

Conditions on Consequence

The logical-consequence relation is evidently truth-preserving: if each of a set of claims is true, then so too are all of its logical consequences. This gives us a very minimal condition on formal systems: given a reading R of the language, it must never be the case that □ $|-_S$ □ if each member of R(□) is true while R(□) is false. Similarly for $|=_S$. It is a relatively straightforward matter to check for satisfaction of this condition, and it is indeed satisfied by all of the standard

[4] See: Etchemendy, *The Concept of Logical Consequence*; Shapiro, "Logical Consequence: Models and Modality;" Sher, *The Bounds of Logic*; McGee, "Two Problems with Tarski's Theory of Consequence;" Blanchette, "Models and Modality."

propositional and first-order formal systems. (For issues about second-order systems' satisfaction of this criterion, see below.)

Presumably, however, we intend a much stronger connection between premises and conclusion than mere truth-preservation when we say that the latter is a logical consequence of the former. We take it, for example, that agents are committed to the logical consequences of the claims they explicitly avow, though of course we do not take them to be committed to all of the truths in virtue of a commitment to one of them. Hence we take it that we can discover unanticipated commitments simply by following out the consequences of our explicit claims. Similarly for theories: the consequences of a theory's assertions are entirely within the realm of claims on the basis of which the theory is to be found adequate or wanting, whether or not one takes theories themselves to be closed under logical consequence. In short, an important feature of logical consequence is that, as we shall say, it *transmits* epistemic and theoretical commitment.

In addition to transmitting commitment, logical consequence would appear to be epistemically inert, in the following sense: The logical consequences of things knowable a priori, or knowable non-empirically or without the aid of intuition, are themselves, respectively, knowable a priori, non-empirically, without the aid of intuition. For any kind K of objects, things knowable without access to objects of kind K pass on this property to their logical consequences. There is a rough sense, then, in which the logical consequences of a claim have no "new content" over and above that had by the original claim. Whether this conception of content can be characterized sufficiently clearly, independently of the relation of consequence, to provide much elucidation here is unclear; for our purposes it will suffice to note that the consequence relation preserves the epistemic categories just noted.

And, finally, there is, it is usually agreed, a certain modal characteristic of logical consequence. The fundamental idea here is that there is a necessary connection between a claim and its logical consequences: if a claim □ is a logical consequence of a set □ of claims, then it is *not possible* for □ to be false while the members of □ are true. Similarly, if an argument is valid, then it is *impossible* for its premises to be true while its conclusion is false.

All of these features of the logical-consequence relation give us conditions that must be met by any reliable deductive or model-theoretic account of consequence. Some of them are more easily formulable and systematically tested than others. We will look here briefly at the criterion given by the last-mentioned condition, namely the modal character of logical consequence.

The relevant criteria of adequacy for the deductive and model-theoretic consequence relations for a system S are that:

(i') If □ $|-_S$ □, then it is impossible for each member of R(□) to be true while R(□) is false; and

(ii') If □ $|=_S$ □, then it is impossible for each member of R(□) to be true while R(□) is false.

As usual, R is an assignment of claims (meeting the usual "rules of thumb") to S's formulas. Only if a formal system S satisfies both (i') and (ii') for every □, □, and R will that system's model-theoretic and deductive consequence relations prove reliable indicators of logical consequence. Satisfaction of (i') is relatively easily established (or refuted): one simply checks each axiom to see that it expresses only necessary truths, and checks each rule of inference to see that it preserves this property. In the case of propositional logic, for example, one simply notes that the axioms (instances of a small handful of forms, like ((A & B) □ A)) express only necessary truths, and that the rule(s) of inference (e.g., *modus ponens*) generate only necessary consequences.

A similar argument *almost* suffices for the usual first-order deductive systems. The only difficult point here concerns the "non-empty universe" assumption built into standard first-order systems. Such formulas as "(□x)x=x" and "(□x)(Fx □ Fx)" are deductive theorems of standard first-order systems. But the claims expressible by these formulas are not uncontroversially necessary truths, and hence not uncontroversially logical consequences of the empty set. If indeed these formulas do express non-necessary truths, then standard first-order systems fail to satisfy (i'), and hence fail to satisfy (i). Aside from these existential formulas, however, it is relatively uncontroversial that standard first-order systems do satisfy the modal requirement (i').[5]

In both the propositional and first-order cases, the completeness theorem enables us to show satisfaction of (ii') via satisfaction of (i'). So standard propositional systems do, and standard first-order systems either do satisfy, or almost satisfy (ii') as well.

We can also establish satisfaction of (ii') directly in certain cases, and it is instructive to see how non-trivial this can be, even in the case of propositional logic. Here, for example, is an argument adapted from one by Richard Cartwright that establishes satisfaction of (ii') by standard systems of classical propositional logic:[6]

Let the language *L* be a set of formulas freely generated from a non-empty set of formulas by the binary operation & and the unary operation N. Let a *valuation* be a function from *L* into {T, F}. Where |= is any relation that takes sets of wffs to wffs, and V is a set of valuations, say that V *induces* |= iff: For every set X of wffs and every wff □, X |= □ iff, for every v □ V, if v(x)=T for every x □ X, then v(□)=T. Roughly speaking, the more valuations V contains, the smaller will be the relation induced by V; where V is empty, the induced relation is universal, and where V contains every valuation, the induced relation will be very small. Say that a valuation v is *Boolean* iff: For all □ and □, v(□) ≠ v(N□), and v(&□□) = T iff v(□)=v(□)=T. Let $|=_B$ be the relation induced by the set of Boolean valuations. Notice that $|=_B$ is the relation of truth-table implication: to say that □ $|=_B$ □ is to say, essentially, that any row of a standard truth-table (treating "&" as conjunction and "N" as negation) that assigns T to every member of □ will assign T to □.

The question of the satisfaction of (ii') by a standard propositional system is the question of whether, for every set □ of wffs and every wff □, if □ $|=_B$ □, then

[5] On the issue of these existential requirements, see Chapter – (Free Logic).

[6] See Cartwright, "Implications and Entailments."

the appropriate necessary connection obtains between the claims made by the members of □ and that made by □. Here we need to know something about the claims expressible by L's formulas. Let these be governed by the usual constraint:

(C) For all formulas □ and □, N□ expresses the negation of what □ expresses, and &□□ expresses the conjunction of what □ and □ express.

Say that a valuation is *admissible* if it represents a possible distribution of truth-values to the formulas, given the constraint (C) on readings. Thus e.g. a valuation assigning T to some wff □ and to N□ is not admissible, since no claim and its negation can both be true. $|=_B$ satisfies (ii') just in case every admissible valuation is Boolean. So suppose v is not Boolean. Then either:

(a) for some wff □, v(□) = v(N□), in which case v is not admissible (since it's not possible for a claim and its negation to be both true or both false);

or

(b) for some wffs □ and □, either v(&□□)= F and v(□) = v(□) = T, in which case v is not admissible (since it's not possible for two claims to be true while their conjunction is false), or v(&□□) = T and either v(□) = F or v(□) = F, in which case v is not admissible (since it's not possible for a conjunction of claims to be true while one conjunct is false).

So every admissible valuation is Boolean. QED.

For systems lacking a completeness theorem, e.g. typical systems of second-order logic, establishment of (ii'), if indeed (ii') holds, must be by some such direct method. The question of the modal adequacy of second-order systems is too large to treat in detail here, but some of the relevant concerns are as follows.

First of all, a complication is that the question of precisely *which* claims a given formula can be taken to express is considerably less clearly answered for second-order systems than it is for first-order and propositional systems. For example, the formula "□X□y(Xy □ y=y)" might, or might not, be taken as an appropriate formalization of the claim that there exists a set of all the self-identical things. If it is taken to be capable of formalizing such a claim, then the system in question will fail (ii'), since the formula is a model-theoretic truth, but the claim is not a necessary truth, and is indeed a falsehood. If on the other hand this reading of the formula is ruled illegitimate, then this particular counterexample to (ii') is not available. Similarly for a large number of potentially problematic model-theoretic truths of second-order logic: on some construals of the expressive power of the language, a variety of such formulas express false claims. Though on such understandings of the language, the model-theoretic consequence relation clearly fails to reliably indicate logical consequence, this does not indict the deductive consequence relation, since again such formal systems lack a completeness theorem. The reliability of the deductive system, i.e. the satisfaction of (i) and of (i'), is of course to be established by looking at the details of particular second-order deductive systems.[7]

For any of the criteria outlined (truth-preservation, topic-neutrality, necessity, etc.), the task of checking the reliability of a particular deductive system is relatively straightforward, since satisfaction of the criteria by the deductive system as a whole can be traced to the satisfaction of these very criteria by the relatively manageable collection of axioms and rules of inference. Checking the reliability of model-theoretic systems, particularly in the absence of a completeness theorem, is often a considerably more difficult matter. Arguments here will sometimes turn on *ad hoc* features of the model-theoretic output of a given system. A nice example of such a feature arises in the case of second-order logic with respect to the continuum hypothesis.[8]

The continuum hypothesis is the hypothesis that there are no sets whose cardinality is larger than that of the natural numbers ($\mathbb{N}$) and smaller than that of the real numbers ($\mathbb{R}$). It is generally agreed (following the work of Gödel and

[7] See Shapiro, *Foundations Without Foundationalism.*

[8] This example is discussed by Etchemendy in *The Concept of Logical Consequence* Ch 8.

Cohen) that the continuum hypothesis (CH) is *independent of* the axioms of ZFC.[9]

The status of the continuum hypothesis makes a difference to model theory. When we ask whether a given formula is true on every model (i.e. is a model-theoretic truth), we are asking whether there *exist* models which falsify that formula. Thus which formulas turn out to be true on every model will depend to a certain extent on what kinds of models - i.e. on what kinds of sets - there are. Because in second-order logic we can define the properties *being of smaller/larger cardinality than* ℕ and *being of smaller/larger cardinality than* ℝ, there are sentences of second-order logic whose status as model-theoretic truths will depend on the disposition of the continuum hypothesis. Specifically, if the continuum hypothesis is true, then the following sentence will be true on every model:

(1) (□X)(X > ℕ □ ℝ □ X),

where "...> N" and "R □..." are abbreviations for the definable properties just noted.

If the continuum hypothesis is false, then there will be models the powerset of whose domain contains sets larger than ℕ but smaller than ℝ, and hence models which falsify (1). In this case, however, the following sentence will be true on every model:

(2) (□X) X > ℕ □ (□X)(X > N & X < ℝ),

with abbreviations as above.

This would seem to pose a problem for the view that the model-theoretic truths of such a language are always logical truths – and hence for the view that model-theoretic consequence in such a system reliably indicates logical consequence. For assuming that the continuum hypothesis really is, as above, independent of the axioms of ZFC, we know that neither the continuum hypothesis nor its negation is a truth of logic. For no truths of logic are independent of ZFC. And if the continuum hypothesis is not a truth of logic, then it is not a truth of logic that every set larger than N is at least as large as R. And

[9] See Chapter 3 of this volume.

since (1) simply says that every set larger than N is at least as large as R, we must conclude that (1) is not a truth of logic. Similarly: if the negation of the continuum hypothesis is not a truth of logic, then it is not a truth of logic that if there are sets larger than N, then there are sets larger than N and smaller than R.[10] So from the fact that the negation of the continuum hypothesis is not a truth of logic, we must conclude that (2) is not a truth of logic either. Hence the problem: Either (1) or (2) is a model-theoretic truth, but neither (1) nor (2) is a truth of logic. So at least one model-theoretic truth is not a truth of logic.

A potential response to this problem is that it simply shows that there is no firm boundary between set theory and logic, hence no firm boundary between set-theoretic truth and logical truth. This may well be so. But it is not of much help for the view that model-theoretic consequence and truth in this system reliably indicate logical consequence and truth. For to defend the reliability of the model-theoretic account, one must hold that either (1) or (2) *is* in fact a logical truth, and hence that either the continuum hypothesis or its negation is a logical truth. But this conflicts with the uncontroversial independence-results which prompt the problem in the first place. If logical truth and set-theoretic truth come to the same thing, then the independence results demonstrate that neither the continuum hypothesis nor its negation is a set-theoretic truth, in which case we see no support for the view that either (1) or (2) is a logical truth. If there are set-theoretic truths which are not logical truths, then the fact (if it is one) that either the continuum hypothesis or its negation is a set-theoretic truth does not support the view that either (1) or (2) is a logical truth. In either case, this example would seem to give us a reason for deeming the usual second-order model-theoretic consequence relation unreliable as an indicator of logical consequence. This does not, of course, provide anything like an indictment of second-order logic in general, and in particular tells us nothing about the reliability of various second-order deductive systems as indicators of logical consequence.

---

[10]Assuming, of course, that it is not a truth of logic that there are no sets larger than N. In this case, though both (1) and (2) will be truths of logic, so too will be the continuum hypothesis.

Different Formal Systems

Not every formal system is designed to reflect the full extent of the logical-consequence relation. Systems of propositional logic, for example, are intended to reflect only a small part of the consequence relation as it applies to the claims expressible in the languages of those systems. Thus one way in which two formal systems can differ over their assessments of logical consequence is that one of the systems can reflect logical consequences not reflected, and not intended to be reflected, by the other. Such differences need not reflect any underlying disagreement about the extension (or nature) of the logical-consequence relation; they can simply be viewed as more or less partial treatments of an agreed-upon relation.

But more robust disagreements are possible as well, disagreements that stem from fundamental disagreements about the nature of the logical-consequence relation itself. As noted above, there are those who hold that the logical consequences of a given claim must have a subject-matter that is in some sense *relevant* to that of the claim itself. On this view, for example, one cannot validly argue from the premises *Jones is wise* and *Jones is not wise* to *Smith is athletic*. Standard propositional and quantified systems of logic count the formalized version of this argument as both deductively and model-theoretically valid, with the result that the relevance-theorist must take those systems to be unreliable indicators of logical consequence. These theorists argue that more reliable indications of consequence and its related logical notions are given by alternative systems of logic, called systems of *relevance* logic. (See chapter 13 of this volume.)

Similarly, systems of *intuitionist* logic are prompted by the perceived unreliability of classical systems. For the intuitionist, it is simply not the case in all domains (for example, when dealing with mathematical existence-assertions) that for each claim □, the corresponding disjunction *either □ or not-□* is always true. On this view, classical logic is wildly unreliable in its assessments of logical consequence. Intuitionist logics are those systems of logic designed to provide reliable indications of logical consequence and related notions as these are understood by the intuitionist. (See chapter 11 of this volume.)

Other disagreements with the classical conception of logical consequence have given rise to yet more alternative systems; see especially chapters 12, 14, 15, 16 of this volume. In all cases, the same principle is at work: a given conception of the pretheoretic relation of logical consequence prompts the construction of a particular kind of formal system, one that will give an accurate, systematic treatment of logical consequence and its related notions.

Analysis of the Relation

We turn, finally and briefly, to the intensional question: What is it that *makes* one claim a logical consequence of others? A response to this question can take one of two forms. The first, dismissive, response is that the relation of logical consequence is primitive and unanalysable, and hence that one cannot reduce the fact of A's following logically from B to any more basic facts about A and B, or to any more basic relationship between them. The second form of response is to explain logical consequence in terms of more fundamental facts about A and B and their relationship to one another.

Looking at the claims expressible by formulas of a particular formal system, one might be tempted to provide an analysis of logical consequence in terms of deducibility in that system, or in terms of truth-preservation across the models of that system. But a moment's reflection will make it clear that no such system-specific analysis of logical consequence can succeed in clarifying what logical consequence consists in, i.e. of what makes it the case that certain claims are logical consequences of others. Deducibility within just *any* system will not do, since there are countless systems, easily definable, which count exactly the wrong things as logical consequences of others. Similarly for model-theoretic consequence. So the attempt to analyze logical consequence via deducibility or model-theoretic consequence must take the analysans here to be deducibility or model-theoretic consequence within a *particular* well-chosen system or kind of system. And the question then arises of what recommends that system or kind of system as an acceptable standard of logical consequence. The attempt to answer *this* question, however, threatens to return us to our original question, that of what makes one claim a logical consequence of others.

An alternative approach is motivated by the fact that logically-valid arguments come in patterns, patterns like Aristotle's syllogistic forms, or the argument-schemes validated by formal systems, or even the natural-language patterns emphasized in teaching critical thinking. Noticing this, it is tempting to define the logical properties and relations in terms of these patterns. Thus for example one might define the logical truths as the instances of patterns each instance of which is true, and similarly for logically-valid arguments.

Whether such a characterization will be extensionally accurate will turn on what counts as a "pattern." The first difficulty here is that patterns themselves are definable only for a given language, and facts about which claims and arguments instantiate the same pattern will vary with the language in question. Consider the argument

Jones and Smith are of the same height

Jones is 6' tall

Therefore, Smith is 6' tall.

Let L1 be a language in which this argument can be formalized as:

h(j) = h(s)

h(j) = a

h(s) = a,

while L2 formalizes it as:

H(j, s)

T(j)

T(s).

Making the obvious parallel assumptions about the other kinds of claims formalizable by these series of formulas in L1 and L2, we see that though each argument formalizable by the L1 series is truth-preserving, this is not the case with the L2 series. So the question of whether our original argument exhibits a pattern each instance of which is truth-preserving, and hence (on the current proposal) the question of whether that argument is valid, will depend on which language we have in mind when characterizing the "pattern." The first problem, then, with the pattern-analysis of the logical relations is that its deliverances will depend in unwanted ways on the language we have chosen to focus on.[11]

There will indeed be languages for which such a pattern-characterization of the logical relations is extensionally accurate, languages in which the sentence-patterns (e.g. "(□ & □) □ □)") each of whose instances is true will turn out to be patterns each of whose instances expresses a logical truth. The clear candidates here are the formal languages of modern logic. Things are less tidy for languages not intentionally designed to have such a result. It is important for the extensional adequacy of the pattern-characterization of logical truth that, e.g., "Smith did it for Jones' uncle" and "Smith did it for Jones' sake" do not count as instantiating the same pattern. And it is important for the attempt to (non-circularly) analyze the logical relations in terms of patterns that the different logical implications had by such pairs of sentences not be appealed to in distinguishing their patterns. In brief, the second difficulty of the pattern-analysis of the logical properties and relations is this: Though we can, for certain specified formal languages, give an extensionally-accurate characterization of the logical properties in terms of truth-preservation across patterns, this is no reason to suppose that the logical properties are *due* to, or explicable in terms of, characteristics of these sentence-patterns. For the formulas of these languages are expressly designed so that they will instantiate the same syntactic patterns when and only when the claims they express have relevantly similar logical properties. We formalize the two English sentences just quoted very differently *because* we recognize that they express claims with very different logical implications; we don't recognize the logical properties on the basis of the

[11] This difficulty remains even when the *claims* are, as above, interpreted sentences. For we presumably want a sentence S to count as a logical truth iff all sentences synonymous with it are as well, and this will not generally be the case on the proposed account.

patterns. And when we turn our attention to natural languages, it is difficult to find a characterization of *patterns* that is plausible, extensionally speaking, without making covert appeal to the very logical properties and relations at issue.[12]

A perhaps more promising approach is the analysis of logical truth as a kind of *analytic* truth. The difficulties of characterizing analyticity itself are legion, but we leave them aside here. The question is whether, granting for the moment the coherence of the notion of analytic truth, we can give an account of logical truth in terms of it. Where analytic truths are, roughly, sentences whose truth is due entirely to matters of meaning (as opposed to matters of "fact"), the logical truths will be those whose truth is due entirely to the meanings of a certain small, select group of terms. These terms, the "logical constants," include the usual "and", "or", "not", "for all", "exists", perhaps "=", and terms definable in terms of these. Thus while "All professors are academics" is arguably an analytic truth in the broad sense, "If all professors are arrogant then all professors are arrogant" falls into the narrower camp of logical truth, since its truth is guaranteed simply by the meanings of its logical constants.

There are at least two difficulties with this approach. The first is, as noted, that it is not entirely clear that the notion of analytic truth can be made sense of. The second is that this characterization of the logical properties and relations seems to appeal, once again, to the very things it is trying to characterize. To say that the meanings of a collection of terms "suffices for" or "guarantees" the truth of a sentence seems to mean little more than that the sentence's truth *follows logically from* facts about those meanings, or that its falsehood would be *logically inconsistent with* those facts, etc. And if this is right, then we cannot without vicious circularity give a characterization of the logical properties and relations in terms of meanings.

This last problem, the circularity of the proposed analysis, would seem likely to pose difficulties for virtually any attempted analysis of the tight circle of inter-defined logical properties and relations. For to give an analysis of logical

[12] See Etchemendy, "Logic as Form."

truth is to say what it is about a given truth that *makes it* a logical truth. Similarly for the notions of validity, logical consequence, inconsistency, and so on. But to say that certain features F of a sentence or claim *make* that claim a logical truth is to say something dangerously close to saying that the sentence's having F *entails* that the sentence is a logical truth. And if we say this, then we have proceeded in a very small circle. Similarly when the analysis is given as a "reduction." We can try to informatively reduce the property of logical truth to a collection of non-logical features F of sentences or claims, holding that to *say* that a claim □ is a logical truth is just to *say* that □ has features F. And we have certainly not in this brief discussion exhausted all of the possible ways of fleshing out such an attempted reduction. But the potential difficulty faced by all such attempts is that of saying precisely how F and the logical relationships in question are related, without making recourse to anything like *entailment* between the two.

Analysis and reduction are typically intimately connected with the logical properties and relationships: we analyze complex notions in terms of simpler ones, or reduce some to others, in part by noting logical connections between the analysans and analysandum. We note facts about *inconsistencies* between affirmations of analysans and denial of analysandum, of *entailments* between claims about one and claims about the other, and so on. If this general pattern is in fact a necessary feature of analysis and reduction, then the logical properties and relations will be analysable in terms of, and reducible to, only other members of the circle of logical properties and relations, and not to any outside it. If so, then we will have to be content with explanations of these notions that consist of making explicit their role in our overall semantic and other cognitive activities, but that do not give simple, informative answers to questions of the form "what *makes* this a logical consequence of that?"

SUGGESTIONS FOR FURTHER READING

Perhaps the most provocative book written in recent years on the topic of logical consequence is Etchemendy's *The Concept of Logical Consequence*, which provides a sustained criticism of the assumption that model-theoretic consequence relations generally provide adequate analyses of logical consequence. Reactions to this criticism can be found in McGee's "Two Problems with Tarski's Theory of Consequence" and Shapiro's "Logical Consequence: Models and Modality." The question of whether model-theoretic consequence relations can guarantee the required modal connection between premises and conclusion is treated Shapiro's paper (just mentioned) and in Blanchette's "Models and Modality."

For discussion of the bearers of the logical relations, and particularly of the difficulties involved in supposing the existence of nonlinguistic propositions, see Quine's "On What There Is" and his *Philosophy of Logic*, esp. Ch. 1. Also see Cartwright's "Propositions" and Strawson's *Introduction to Logical Theory*, esp. Ch. 1.

The classic criticism of the notion of analytic truth and related notions, together with an influential treatment of logical truth, can be found in Quine's "Two Dogmas of Empiricism." A response is found in Strawson's "Propositions, Concepts, and Logical Truths."

A number of useful papers on related topics can be found in Hughes' *Philosophical Companion to First-Order Logic*. A very readable discussion of many of these issues can be found in Haack's *Philosophy of Logics*.

REFERENCES

Blanchette, Patricia: "Models and Modality," *Synthese*, forthcoming.

Cartwright, Richard: "Implications and Entailments," in Cartwright's *Philosophical Essays* 237-256.

Cartwright, Richard: "Propositions," *Analytical Philosophy*, 1st Series; R. J. Butler (ed), Oxford: Blackwell 1962, pp 81-103. Reprinted in Cartwright's *Philosophical Essays* 33-53.

Cartwright, Richard: *Philosophical Essays,* MIT Press, Cambridge and London 1987.

Etchemendy, John: "The Doctrine of Logic as Form," *Linguistics and Philosophy* 6, 319-334.

Etchemendy, John: *The Concept of Logical Consequence*, Harvard University Press, Cambridge 1990. Reprinted by CSLI Publications, Stanford 1999.

Frege, Gottlob: *Grundgesetze der Arithmetik* (Vol I: 1893; Vol II: 1903), partially translated by M. Furth as *The Basic Laws of Arithmetic*, University of California Press 1964.

Frege, Gottlob: *Philosophical and Mathematical Correspondence*, Gabriel, Hermes, Kambartel, Thiel, Veraart, McGuinness (eds), University of Chicago Press 1980.

Haack, Susan: *Philosophy of Logics*, Cambridge University Press 1978.

Hughes, R.I.G.: *A Philosophical Companion to First-Order Logic*, Hackett Publishing Company, Indianapolis/Cambridge 1993.

McGee, Vann: "Two Problems with Tarski's Theory of Consequence," *Proceedings of the Aristotelian Society* 92 (1992) 273-292.

McGee, Vann: Review of Etchemendy's *The Concept of Logical Consequence*, *The Journal of Symbolic Logic* 57 (1992) 254-55.

Quine, W.V.O.: "On What There Is," *Review of Metaphysics* 2 (1948) 21-38; reprinted in *From a Logical Point of View*.

Quine, W.V.O.: "Two Dogmas of Empiricism," *Philosophical Review* 60 (1951), 20-43; reprinted in *From a Logical Point of View*

Quine, W. V. O.: *From a Logical Point of View*, Harvard University Press 1953.

Quine, W. V. O.: *Philosophy of Logic*, Prentice Hall, Englewood 1970.

Shapiro, Stewart: "Logical Consequence: Models and Modality" *The Philosophy of Mathematics Today*, M. Schirn (ed) Clarendon Press, Oxford.

Sher, Gila: *The Bounds of Logic*, MIT Press, Cambridge 1991

Sher, Gila "Did Tarski Commit Tarski's Fallacy?" *Journal of Symbolic Logic* 61 (1996) 653-686.

Strawson, P.F.: *Introduction to Logical Theory*, Methuen & Co, London 1952.

Strawson, P. F.: "Propositions, Concepts, and Logical Truths," *The Philosophical Quarterly* 7 (1957). Reprinted in Strawson's *Logico-Linguistic Papers* 116-129.