4

The Science of Logic: An Overview

1. INTRODUCTION[1]

Our discussions in chapter 3 of the nature, scope, and modes of human knowledge have helped prepare the way for an overview, in the present chapter, of the science of logic. Here we shall give an account, in deliberately general terms, both of *what* we know of the subject matter of logic and of *how* we know it.

Logic, as a science, has much in common with the other special sciences — mathematics, physics, chemistry, and the rest. Like each of them it is no mere collection of known propositions, but is rather an *organized* body of knowledge with its own highly general principles and laws. And like each of them it has its own *methods* of inquiry, its own distinctive ways of expanding knowledge.

Yet logic also has a special status among the sciences. In the first place, it is the most general science insofar as all propositions whatsoever fall within its compass, whereas each of the other sciences treats of relatively restricted sets of propositions — propositions about numbers, about material particles, about chemical properties, and so on. This fact has led some logicians to claim, somewhat paradoxically, that logic has no subject matter at all and that it is the science of pure form. (We shall try to put this claim into perspective later in this chapter.) In the second place, logic shares with mathematics the distinction of relying on methods which are purely a priori, whereas each of the remaining sciences relies on methods which are largely empirical.[2] This fact has led some thinkers to claim that arithmetic is but an extension of logic. Whether or not there is any line of demarcation between the two and, if so, where it should be drawn is still an issue much in dispute within the philosophy of mathematics. We will make no further comment upon it. Whatever the outcome, it would generally be agreed — except by Radical Empiricists like Mill — that the methods of logic (and probably those of mathematics) are wholly ratiocinative and hence a priori.

1. The material in this chapter provides a natural bridge between chapters 3 and 5; however, it is both complex and condensed. Teachers may wish to make judicious selections from among its six sections.

2. But not wholly. Each of the so-called "empirical sciences" makes use as well of the a priori principles and findings of logic (and in many cases, of mathematics). This is not surprising in the light of our definition of "empirical". For, it will be remembered, to say that a proposition is knowable empirically is to say that it is knowable only experientially, while to say that a proposition is knowable only experientially is just to say that the only way in which it can be known is by direct appeal to experience *or by inference therefrom.* And this is where the a priori science of logic (and, in many cases, of mathematics) plays a role. For an inference that is made from a proposition which is known by appeal to experience will be known to be a *valid* inference if and only if it accords with the a priori principles and findings of logic.

We shall start our overview of the science of logic by saying something more about these methods.

Since there are two main ways of acquiring ratiocinative knowledge — by analysis and by inference therefrom — it follows that, to the extent that our knowledge of logical propositions is a priori, it too is gained by analysis or inference. To be sure, it is possible to acquire knowledge of some necessary truths, including some truths of logic, experientially. But experience, we have seen (following Kant), cannot give us knowledge of the modal status of these propositions. Nor does it offer us any surety as a method for systematically expanding our knowledge of the subject matter of logic. On the contrary, such knowledge as it gives us arises, as it were, adventitiously. The a priori methods of analysis and inference offer us the best prospect for building a *science* of logic.

2. THE METHOD OF ANALYSIS

The method of analysis has always been employed in the sciences of mathematics and logic — and, for that matter, philosophy. Euclid employed it when, in his *Elements,* he analyzed the concepts of being a point, being a straight line, triangle, etc. In this way he laid the foundations of geometry. Plato employed it when, in his *Theatetus,* he analyzed the concept of knowledge, and again, in his *Republic,* when he analyzed the concept of justice. In this way, Plato laid the foundations of two important subdisciplines of philosophy: epistemology (i.e., theory of knowledge) and philosophy of politics, respectively. Aristotle employed it when, in his *Organon,* he analyzed syllogistic reasoning and the modal concept of necessity; the Megarian logicians employed it when, in miscellaneous inquiries, they analyzed conditional propositions. Between them, Aristotle and the Megarians laid the foundations of logic.[3] In all these as well as other fields, analysis is still fundamental.

But what does analysis, in philosophy and logic, consist of? What are its objects? And what are its results?

Analysis, in general, as we pointed out in chapter 2, consists in the examination of a complex item of some sort with a view to determining what constituents make it up and how they are related to one another. This is evident in fields such as chemistry, grammar, and the like. Philosophical analysis differs from these and other analytical enterprises primarily in the nature of the complex items it examines. In chemistry one analyzes chemical compounds. In grammar one analyzes sentences. What does one analyze in philosophy, in general, and in logic, in particular?

The objects of philosophical analysis

That propositions and concepts, insofar as they are complex, are *among* the items which we analyze when doing philosophy, was noted when — in chapter 2 — we developed the theory that propositions are truth-valued combinations of concepts. G. E. Moore, in his classic answer to the question, wrote as if they were the *only* objects of analysis,

> In my usage [of the term 'analysis'] both *analysandum* and *analysans* must be concepts or propositions.[4]

3. For more on the contributions of the Megarians, see W. and M. Kneale, *The Development of Logic,* Oxford, Clarendon Press, 1962, chapter 3, esp. pp. 113–138.

4. See "A Reply to My Critics", in *The Philosophy of G.E. Moore,* ed. P.A. Schilpp, La Salle, Open Court, 1968, vol. 2, p. 664. By the term "analysandum" we mean the expression of the object of analysis; by the term "analysans" we mean the expression of the result of analysis. Compare the parallel accounts of the terms "definiendum" and "definiens" given in footnote 18, p. 26, chapter 1. Definitions are (usually) of verbal

One might object that Moore's answer is too restrictive since it seems to preclude our giving analyses of *arguments* which are surely among the prime objects of analysis in logic. Then again, it seems to preclude our giving analyses of *questions* and *commands,* items to which philosophers, especially recently, have devoted a good deal of close analytical scrutiny.

But these are quibbles. An argument, from the point of view of analysis, is just a pair of proposition-sets standing in relation to one another as premises to conclusion. And questions and commands, although they are not themselves propositions or concepts, are analyzable (just as propositions are) into conceptual constituents — albeit ones which stand in non-truth-valued types of combination. In any case, so far as our present interest is concerned — that of describing the method of analysis as it applies within the science of logic — Moore's answer is entirely adequate. We can afford to ignore questions, commands, requests, prayers, and the like. For the preoccupation of the science of logic is, as we might have expected, a restricted one: it is concerned almost solely[5] with the analysis of propositions, of concepts (insofar as they feature in propositions) — and, as an aside, of arguments (insofar as propositions feature in them).

Three levels of analysis

There seem to be three main forms which analysis can take in logic:

1. that in which the analysandum is a proposition while the analysans features, as the constituents of that proposition, a simpler proposition (or set of propositions) together with a simpler concept (or set of concepts);
2. that in which the analysandum is a proposition while the analysans features, as the constituents of that proposition, a set of concepts; and
3. that in which analysandum is a concept while the analysans features, as the constituents of that concept, a set of simpler concepts.

We shall consider one or two examples of each form of analysis.

Examples of form 1:

Consider, first, the proposition

(4.1) Either it is necessarily true that sisters are female or it is not.

Plainly *(4.1)* is admissible as an object of analysis. For it is a complex which has other *propositions* among its constituents, viz.,

(4.2) It is necessarily true that sisters are female

and

(4.3) It is not necessarily true that sisters are female.

expressions; analyses, as Moore would insist, are never of verbal expressions but always and only of what verbal expressions express, viz., concepts or propositions.

5. We say "almost solely" in order to allow for the fact that some philosophers have tried to construct what they call logics of nonpropositional kinds. See, for instance: Charles L. Hamblin "Questions", *Australasian Journal of Philosophy,* vol. 36 (1958), pp. 159–168; and David Harrah, "A Logic of Questions and Answers", *Philosophy of Science,* vol. 28 (1961), pp 40–46. Attempts have also been made to construct logics of imperatives.

Any proposition which has other propositions among its constituents is what we call a *compound proposition*. Hence, *(4.2)* is a compound proposition. By way of contrast,

(4.4) Sisters are female

is *not* a compound proposition since it does not have any other propositions among its constituents. Such a proposition we call a *simple proposition*. Note that the propositions *(4.2)* and *(4.3)*, which are constituents of the compound proposition *(4.1)*, are themselves compound propositions. Though they are *simpler* than *(4.1)* they are nevertheless not *simple* propositions since each has the even simpler proposition *(4.4)* as one of its constituents. In short, *(4.1)* may be regarded as having *(4.2)* and *(4.3)* among its constituents even though each of the latter is itself susceptible to being analyzed into still simpler constituents. Another constituent of *(4.1)* is the concept of *disjunction*, i.e., the concept of that relation which holds between two (or more) propositions in all those possible worlds (if any) in which at least one of them is true or between two (or more) concepts in all those possible worlds (if any) in which at least one of them has application. In short, the constituents of *(4.1)* are the two propositions, *(4.2)* and *(4.3)*, plus the concept of disjunction; and, within *(4.1)* these constituents stand in propositional combination.

As a second example of analysis of the first form consider the proposition

(4.2) It is necessarily true that sisters are female.

We have already noted that *(4.2)* features as a propositional constituent in the analysans of *(4.1)*. Yet it, too, may be analyzed since, it, too, is a complex which has simpler constituents. One of these constituents, we have seen, is the simple proposition *(4.4)*. Another is the concept of necessary truth. And within *(4.2)* these two constituents — one a proposition, the other a concept — stand in propositional combination insofar as the property of necessary truth is predicated of, or ascribed to, the proposition *(4.4)*.

Before proceeding to examples of analyses of the second and third forms, we would do well to note a common feature of analyses of the first form. In each case, the analysandum is shown by analysis to have, among its constituents, propositions which, though they themselves are susceptible to analysis, are nevertheless *at this level of analysis* left *unanalyzed*. For this reason we shall refer to that part of the science of logic which deals with propositions at this level of analysis as *The Logic of Unanalyzed Propositions*.

Example of Form 2:

(4.4) Sisters are female.

As already noted, *(4.4)* features as a propositional constituent in the analysans of *(4.2)*. Yet it, too, may be analyzed. Although it is a simple proposition, with no other propositions among its constituents, it is nevertheless complex insofar as it has concepts as its constituents, viz., the concepts of being a sister and of being female. And, within *(4.4)*, these two constituent concepts stand in propositional combination insofar as the property, of which being female is the concept, is predicated of, or ascribed to, the items to which the concept of being a sister applies.

A common feature of all analyses of the second form is that each has as its analysandum a proposition which is shown by analysis to have, among its constituents, concepts which, though they themselves are susceptible to analysis, are nevertheless *at this level of analysis* left *unanalyzed*. For this reason we shall refer to that part of the science of logic which deals with propositions at this level of analysis as *The Logic of Unanalyzed Concepts* (often known as The Logic of Predicates).

Example of Form 3:

(4.5) being a sister

Although the concept of being a sister features as a conceptual constituent in the analysans of *(4.4)*, it itself is subject to analysis. For it, too, is complex. Among the simpler concepts which are its constituents are the concepts of being female and of being a sibling. Another constituent is the concept of *conjunction,* i.e., the concept of that relation which holds between two (or more) propositions in all those possible worlds (if any) in which both of them are true, or — as in the present case — between two (or more) concepts in all those possible worlds (if any) in which both of them have application to the same item. In short, the constituents of *(4.5)* are the concepts of being female, being a sibling, and conjunction.

A common feature of all analyses of the third form is that each has as its analysandum a concept which is shown to be analyzable into simpler (though not necessarily simple) concepts. For this reason we shall refer to that part of the science of logic which deals with concepts at this level of analysis as *The Logic of Analyzed Concepts.*

The idea of a complete analysis

Now it is clear, upon review, that each successive form of analysis considered is, in a fairly precise sense of the word, a *deeper* analysis than that of the preceding form. That which features as a constituent in the analysans corresponding to a given analysandum may in turn be analyzed into even simpler constituents. This fact in no way precludes the possibility of our giving an analysis of a certain form — or, as we have otherwise put it, "at a certain level" — which is a *complete analysis relative to that level.* But it has led some philosophers to entertain the ideal of an *absolutely complete analysis* — the ideal of a type of analysis which would involve breaking down a proposition or concept into constituents which are *ultimately simple* insofar as they do not themselves have any simpler constituents and so do not admit of any further analysis. The early Wittgenstein thought in this way. He argued, in his *Tractatus Logico-Philosophicus,* that if any propositions or concepts are to be determinate, then complex propositions must be analyzable into ultimately simple propositions and these in turn must be analyzable into those ultimately simple constituents for which he reserved the word "names".[6] And Moore, although he seems not to have committed himself to the extremes of Wittgenstein's position, thought that there are *in fact* some concepts which are ultimately simple and unanalyzable. In his *Principia Ethica,* for example, he came to this conclusion about the concept of goodness.[7] Fortunately, we do not have to settle here the question whether there are any ultimate simples of analysis, let alone whether there *must* be such. It suffices for us to learn the lesson which both these philosophers and their contemporary, Bertrand Russell, never tired of teaching, viz., that a proposition or concept, the grammar of whose expression appears simple, may well turn out, on analysis, to be logically complex. Indeed Russell's analysis[8] of the seemingly simple proposition

6. Ludwig Wittgenstein, *Tractatus Logico-Philosophicus,* trans. D.F. Pears and B.F. McGuinness, London, Routledge & Kegan Paul, 1961; especially theses 3.23, 3.25, 3.201, and 3.202. See also his *Notebooks 1914–16,* ed. G.H. von Wright and G.E.M. Anscombe, trans. G.E.M. Anscombe, New York, Harper & Brothers, 1961, p. 46.

7. G.E. Moore, *Principia Ethica,* Cambridge, Cambridge University Press, 1962 (copyright 1903).

8. B. Russell, "On Denoting", *Mind,* vol. 14 (1905), pp. 479–493.

(4.6) The present King of France is bald

into the conjunction of three simpler propositions, viz.,

(4.7) There is some item which is King of France,

(4.8) At most one item is King of France,

and

(4.9) Any item which is King of France is bald,

is widely, and deservedly, regarded as one of the classics of analytical philosophy. To be sure, each of these three propositions which feature in the analysans of *(4.6)* is itself susceptible to analysis at a *deeper* level. But at the level at which Russell's analysis is given, (viz., the Logic of Analyzed Propositions), it counts as a *complete* analysis.

The need for a further kind of analysis

Analysis, at any of the three levels distinguished, and whether complete or partial, can make it possible for us to obtain knowledge a priori of the *relations* between the propositions or concepts which feature as the objects of analysis and other propositions or concepts.

In the first place, it yields knowledge of the relation between the analysandum and the analysans. For, if the analysis is sound and complete, it shows that the analysandum is *equivalent* to the analysans. And, if the analysis is sound but only partial or incomplete, it shows that the analysandum *implies* the analysans. For instance, the analysis offered above of the concept of being a sister is a complete analysis and hence, if sound, shows that this concept is *equivalent* to the concept of being both female and a sibling. By way of comparison, Kant's analysis of the concept of body, discussed in chapter 3, purports to be only a partial analysis and hence, if sound, shows only that the concept of body *implies* the concept of extension. The knowledge acquired in each case is, of course, acquired without any need of appeal to experience. It is acquired by virtue of, as Kant would put it, "reason's own resources", and hence is acquired a priori.

In the second place, the types of analysis so far considered may prepare the ground for discovering previously undiscovered relations between the proposition or concept featured in the analysandum and some proposition or concept which does *not* feature in its analysans. For instance, Russell's analysis of the proposition *(4.6)*, that the present King of France is bald, gives us grounds for inferring that this proposition is inconsistent with any proposition which is inconsistent with any of the propositions featured in its analysans. For example, it would be inconsistent with the proposition that there are at present several Kings of France.

Now it is all very well to learn, by analysis, that a proposition, P_1, is equivalent to or implies a proposition, P_2; or, again, that a concept, C_1, is equivalent to or implies a concept, C_2. But this in itself does not tell us whether P_1 is a true proposition; nor does it tell us whether concept C_1 has application in this or any other possible world. Analysis which merely tells us what are the constituents of a proposition can tell us that *(4.1)* is equivalent to the disjunction of *(4.2)* and its negation, viz., *(4.3)*, but — by itself — it cannot tell us anything about the truth-value or modal status of *(4.1)*. Again, analysis which merely tells us what are the constituents of a concept can tell us that the concept of a greatest prime number is equivalent to the concept of a number which has no factors except itself and one, and which has no successor which is a prime, but — by itself — it cannot tell us whether the concept of a greatest prime number has application in this or any other possible world.

Yet have we not spoken hitherto of analysis as a means whereby one might in certain cases ascertain the *truth-value* of a proposition — in an a priori manner — *by analyzing that very proposition?* How can this be possible?

The answer is that if, at any of the three levels of analysis distinguished — that of the logic of Unanalyzed Propositions, that of the Logic of Unanalyzed Concepts, or that of the Logic of Analyzed Concepts — we wish to ascertain the truth-value of the proposition analyzed or the applicability of the concept analyzed, we must supplement analysis into constituents by what is commonly called "truth-condition analysis", or what we prefer to call "possible-worlds analysis".[9] It is only when "analysis" is so understood that we can make sense of the notion of analytically determined truths.

Possible-worlds analysis

By "possible-worlds analysis" we mean the investigation of the conditions under which a proposition is true or the conditions under which a concept has application — where, by "conditions", we do not mean conditions in the actual world but mean, rather, conditions in any possible world. In other words, possible-worlds analysis is the investigation which sets out to determine whether a proposition is true or a concept has application in all, in none, or in some but not all possible worlds. It is the sort of analysis which *can* tell us, for instance, that the concept of the greatest prime number has no possible application — that it is (as we put it in chapter 2) a necessarily non-applicable concept. And it is the sort of analysis which *can* tell us, for instance, that proposition *(4.1)* is true in this as well as all other possible worlds — that it is a necessarily true proposition.

That analysis of propositions and concepts should — in the last resort — involve reference to possible worlds is only to be expected, since (as we argued in chapter 2) the explication of concepts involves reference to sets of possible worlds, and propositions just *are* truth-valued combinations of concepts. But quite independently of that, the link with possible worlds remains. It has been made, implicitly at least, and explicitly in many cases, ever since the dawn of analytical philosophy, even though many analysts would not have thought of so describing it.

Consider, once more, Kant's treatment of the proposition

(3.13) All bodies are extended.

Kant's analysis of *(3.13)* is, in effect, an amalgam of analysis into constituents and of possible-worlds analysis. He calls *(3.13)* an analytic judgment on the grounds (a) that it has among its constituents the concepts of body and of extension, and (b) that the concept of extension is "bound up with" the concept of body. Elsewhere, in characterizing analytic propositions more generally, he speaks of these as propositions in which the concept of the subject covertly "contains" the concept of the predicate. Now it is clear that what Kant means by his "binding" and "containment" metaphors is simply this: that in the case of an analytic proposition it is not *possible* for the concept of the subject to apply without the concept of the predicate also applying. In particular, he may be understood as saying of *(3.13)* that it is not possible for the concept of body to have application to a particular item unless the concept of extension also has application to the same item. But, as we have already seen — in our own earlier analyses of the modal concepts of possibility, necessity, etc. — this is just to say that there is no *possible world* in which the concept of the body applies to an item when the concept of extension does not;[10] which is to say that the concept of body *implies* the concept of

9. We prefer talk of "possible-worlds analysis" not only because it links the notion of analysis firmly with much else that we have said in this book, but also because it is less restrictive than talk of "truth-condition analysis". The latter sort of talk is appropriate only in connection with propositions, not in connection with concepts. Concepts have *applicability* conditions but not *truth* conditions.

10. Note that, on this account, to say — with Kant — that concept C_1 "contains" concept C_2 is to say that

extension; which is to say that it is *necessarily true* that if something is a body then it is extended. Thus it is that analysis into constituents, combined with possible-worlds analysis, can lead to knowledge of the contingency or noncontingency of a proposition and sometimes also of its truth-value.

Significantly, the analysis, at level 3, of a concept such as those of being a body or of being a sister can — when supplemented by analyses of other concepts — yield knowledge of the modal status and truth-value not only of the proposition within whose analysans that concept, at level 2, features as a constituent, but can also sometimes help us acquire knowledge of the modal status and truth-value of propositions at level 1. To illustrate, let us revert once more to the series of propositions *(4.1)*, *(4.2)*, *(4.3)*, and *(4.4)* within each of which the concept *(4.5)*, of being a sister, features as a constituent.

In analyzing

(4.5) being a sister

we are, in effect, attempting to say *what it is* to be a sister by determining in which possible worlds *(4.5)* has application. And our answer is that *(4.5)* has application in a certain set of possible worlds, viz., that which may summarily be described by saying that it is the set in which both the simpler concepts, of being female and of being a sibling, have application to the same item. This means, of course, that the analysans is *equivalent* to the analysandum. For, on the account we gave (in chapter 2) of the equivalence of concepts, concepts are equivalent if and only if in any possible world in which one has application to a given item, the other has application to the same item. It means, too, that the concept of being a sister *implies* each of the concepts featured in the analysans, separately and jointly: it implies being female; it implies being a sibling; and it implies being both female and a sibling. And that, in turn, means that it is *necessarily true* that if something is a sister then it is female; that it is *necessarily true* that if something is a sister then it is a sibling; and that it is *necessarily true* that if something is a sister then it is both female and a sibling.

It is evident that one of the consequences of our determining the applicability conditions of the concept *(4.5)* is that we have thereby also determined both the truth-value and modal status of the proposition

(4.4) Sisters are female

within which *(4.5)* occurs as a constituent. For to conclude, as we just did, that it is necessarily true that if something is a sister then it is female, *just is* to conclude that *(4.4)* is both true and necessarily so. Thus, by a process of reasoning analogous to Kant's, we have established that *(4.4)* is analytically and *a fortiori* necessarily true.[11]

Once this has been established it is, of course, a relatively straightforward matter — involving only an additional understanding, through analysis, of the concept of necessity — to determine both the modal status and the truth-value of

---

the set of possible worlds in which C_2 has application includes or contains the set of possible worlds in which C_1 has application — not, as might have been supposed, the other way around. This curious 'reverse' mapping of propositions onto sets or classes is not unique to this one occasion. It is a universal phenomenon. Consider the proposition expressed by the sentence "A thing's being both red and square implies that it is red." The first class just referred to is the class of red and square things, and the second, the class of red things. Which of these two classes 'contains' the other? Clearly, the *latter* contains the *former*.

11. Recall from chapter 3, p. 146, that an analytically true proposition is one which is known a priori *by analysis* to be necessarily true.

(4.2) It is necessarily true that sisters are female.

For *(4.2)* simply asserts of *(4.4)* that *(4.4)* has the property of necessary truth which our analysis of *(4.4)* shows it to have. In other words, since analysis of *(4.2)* shows it to predicate of *(4.4)* the very property which *(4.4)* has been shown, by analysis, to have, *(4.2)*, like *(4.4)*, can be known to be true analytically; and hence *(4.2)*, like *(4.4)*, is necessarily true.[12]

It then becomes a straightforward matter — involving only the additional understanding, through analysis, of the concept of negation — to determine both the modal status and the truth-value of

(4.3) It is not necessarily true that sisters are female.

For since *(4.3)* is the negation of *(4.2)*, it will be false in all those possible worlds in which *(4.2)* is true. And since *(4.2)* has been shown to be true in *all* possible worlds, we can conclude that *(4.3)* is false in all possible worlds, i.e., that it is necessarily false.

Moreover, once we have established, by reasoning of the kind just displayed, that *(4.2)* is true, and indeed that it is necessarily true, it is but a small step — involving only an additional understanding through analysis, of the concept of disjunction — to determine the truth-value, and indeed the modal status, of

(4.1) Either it is necessarily true that sisters are female or it is not.

Degrees of analytical knowledge

Exactly how much we can come to know about *(4.1)* is a function of how much analytical knowledge of its constituents we take into account. If we take into account merely the fact that *(4.2)* is true, and leave out of account the fact that analysis can show it to be necessarily true, we can — by invoking the analysis of the concept of disjunction — show *(4.1)* to be true. For the concept of disjunction is the concept of a relation which holds between propositions (and derivatively between concepts) when at least one of those propositions is true (or, in the case of concepts, when at least one of those concepts has application to a given item). Thus, since *(4.1)* involves the disjunction of two propositions, *(4.2)* and *(4.3)*, of which we know *(4.2)* to be true, *(4.1)* must itself be true. Clearly, then, knowledge of the analysis of the concept of disjunction, taken together with knowledge of the truth of one of the propositions disjoined in *(4.1)*, suffice to give us knowledge of the *truth-value* of *(4.1)*. But it tells us nothing of its modal status. We can, however, determine the modal status of *(4.1)* in either of two ways. (a) We can tell that *(4.1)* is *necessarily* true, if we take into account the additional fact — already ascertained by analysis — that one of the disjuncts in *(4.1)* is the negation of the other. Thus, since *(4.3)* is the negation of *(4.2)*, and hence (as we have already seen) is false in all those possible worlds in which *(4.2)* is true (and *ipso facto* is true in all those possible worlds in which *(4.2)* is false), these two disjuncts exhaust the set of all possible worlds in the sense that in all possible worlds one or other of them is true. But this is just to say that in all possible worlds either *(4.2)* or *(4.3)* is true; which is to say that *(4.1)* — which asserts that

12. The reasoning can be generalized. Given any proposition, P_1, whatever, if analysis of P_1 shows P_1 to be necessarily true, then any proposition, P_2, which asserts that P_1 is necessarily true, is itself knowable by analysis and hence is necessarily true; likewise for any proposition, P_3, which asserts of P_2 that it is necessarily true; and so on *ad infinitum*. Here we have proof of the thesis $\Box P \rightarrow \Box\Box P$ — a thesis which, taken together with the converse thesis $\Box\Box P \rightarrow \Box P$, yields one of the so-called "reduction laws", viz., $\Box P \leftrightarrow \Box\Box P$, of the systems S4 and S5 of modal logic. (See pp. 220–224, for more on reduction laws.)

either *(4.2)* or *(4.3)* is true — is necessarily true. (b) Again, we can tell that *(4.1)* is *necessarily* true, if we take into account the fact — already ascertained by analysis — that one of its disjuncts, viz., *(4.2)*, is necessarily true. For the analysis of the concept of disjunction shows that the disjunction of any two propositions is true in all those possible worlds in which either disjunct is true. If, therefore, as we have already ascertained by analysis, *(4.2)* is not only a disjunct in *(4.1)* but also is necessarily true, then *(4.1)* will be true in all those possible worlds in which *(4.2)* is true, viz., in all possible worlds. Once more, then, *(4.1)* may be shown, by analysis, to be necessarily true.

Plainly, the analysis of a concept at the deepest level, viz., level 3, may, when supplemented by possible-worlds analyses, at levels 2 and 1, of other concepts, yield knowledge a priori of the modal status of a wide range of propositions and of the truth-value of some of them. At each level one has only to incorporate the results of conceptual analysis at the deeper level in order to acquire this knowledge. (Perhaps it is for this reason that many philosophers refer to *all* analysis as *conceptual analysis* even when the ostensible analysandum is not a concept but a proposition.)

By the same token, if we *neglect* to take into account knowledge acquired by analysis at a deeper level, we can preclude ourselves from acquiring certain sorts of knowledge about the propositions analyzed. This holds even *within* a given level. For instance, within level 1 — the level which leaves certain constituent propositions unanalyzed — if we analyze *(4.1)* merely to the point of recognizing it as the disjunction of two simpler propositions, we cannot thereby determine its modal status and hence cannot determine its truth-value either. If, however, we analyze it more deeply as involving the disjunction of two propositions one of which is the negation of the other, then we *can* — by analyzing the concepts of disjunction and negation — determine both its truth-value and its modal status. Likewise, if we analyze *(4.2)* merely to the point of recognizing it as a proposition which attributes the property of being necessarily true to the simpler proposition *(4.4)*, we can determine *(4.2)*'s modal status but not its truth-value: we can tell that it is noncontingent but not whether it is noncontingently true (necessarily true) or noncontingently false (necessarily false).[13] If, however, we analyze *(4.2)* more deeply, as involving the attribution of necessary truth to a proposition which itself is shown, by analysis, to be necessarily true, then we can tell its truth-value as well as its modal status: we can tell that it is necessarily true, not just that it is noncontingent. And again, if we analyze *(4.4)* in turn merely to the point of recognizing it as a proposition which attributes the property of being female to those items which have the property of being sisters, we cannot determine either its truth-value or its modal status. If, however, we analyze *(4.4)* more deeply, so as to reveal that the set of possible worlds in which the concept of being a sister has application to a given item is included within the set of possible worlds in which the concept of being female has application to the same item, we can then tell that *(4.4)* is — as we saw before — not just true but necessarily so.

The point we are making may be put more generally by saying that although analysis at the level of the Logic of Unanalyzed Propositions can reveal the modal status, or the modal status and the truth-value of certain propositions, these propositions are a proper subset of those whose modal status, or modal status and truth-value, can be determined by analysis at the level of the Logic of Unanalyzed Concepts (the Logic of Predicates); and the latter propositions are a proper subset of those whose modal status, or modal status and truth-value, can be determined by analysis at the level of the Logic of Analyzed Concepts.

The method of analysis, it is clear, is capable of yielding knowledge of the modal status of propositions, and in the case of noncontingent propositions often of the truth-value as well, wholly a priori, i.e., without need of any recourse to experience of how things stand in the actual world.

13. We show this in chapter 6 when we argue that possible-worlds analyses show that *all* attributions of modal status to propositions are noncontingent, i.e., either necessarily true or necessarily false. See pp. 333–36.

3. THE PARADOX OF ANALYSIS

Despite what we have just said, it has seemed to some philosophers that analysis can *never* yield knowledge of anything that we do not already know. Indeed it is very easy to be beguiled by the so-called "Paradox of Analysis" into concluding that the method of analysis is either superfluous or productive of conceptual error.

Moore's problem

The paradox received its classic formulation in C. H. Langford's essay, "The Notion of Analysis in Moore's Philosophy":[14]

> Let us call what is to be analyzed the analysandum, and let us call that which does the analyzing the analysans. The analysis then states an appropriate relation of equivalence between the analysandum and the analysans. And the paradox of analysis is to the effect that, if the verbal expression representing the analysandum has the same meaning as the verbal expression representing the analysans, the analysis states a bare identity and is trivial; but if the two verbal expressions do not have the same meaning, the analysis is incorrect.

For instance — changing examples to one discussed by Moore — the analysis of the concept

(4.10) being a brother

may be stated in some such way as the following:

(i) "The concept of being a brother is identical with the concept of being a male sibling";

(ii) "The propositional function 'x is a brother' is identical with the propositional function 'x is a male sibling' ";

or

(iii) "To say that a person is a brother is the same thing as to say that that person is a male sibling."[15]

But no matter in which way one states the analysis — no matter what verbal expression one gives of the analysandum and the analysans — the paradox presents itself. If the proposition expressed by any of the sentences, (i), (ii), or (iii), is true, then it seems that the very same proposition may be expressed by saying

(iv) "The concept of being a brother is identical with the concept of being a brother."

That is to say, if the analysis is correct then the proposition expressed by any of (i), (ii), or (iii) is identical with the proposition expressed by (iv). But (iv) — it is said — is "trivial" in the sense that it gives no information. Hence if any analysis of *(4.10)* is correct, it is trivial. And equally, of course, if it is not trivial then it is not correct.

14. Published in *The Philosophy of G.E. Moore,* pp. 321–41, esp. p. 323.

15. These three alternative ways of expressing the analysis are suggested by Moore in his "Reply to My Critics", *The Philosophy of G.E. Moore,* pp. 535–687n. Note that mode (ii) makes it clear that Moore, too, regards a concept as what is expressed by certain sorts of propositional function (i.e., "open sentence" as we earlier put it in chapter 2).

A Moorean solution

The paradox, it has been pointed out,[16] is a special case of the paradox of identity. How can *any* statement of identity, if true, be informative and non-trivial? Since the number 9 is identical with the number 3^2, how can the equation "$9 = 3^2$" be more informative than the trivial equation "$9 = 9$"?

A good many suggestions have been made as to how the paradox can be avoided.[17] Many of these are too lengthy for us to consider here. In any case our own preference is to follow up on some suggestions made by Moore in his "Reply" — suggestions which he himself did not quite see how to carry through:

> I think that, in order to explain the fact that, even if "To be a brother is the same thing as to be a male sibling" is true, yet nevertheless this statement is *not* the same as the statement "To be a brother is to be a brother", one *must* suppose that both statements are in *some* sense about the expressions used as well as about the concept of being a brother. But in *what* sense they are about the expressions used I cannot see clearly; and therefore I cannot give any clear solution to the puzzle. The two plain facts about the matter which it seems to me one must hold fast to are these: That if in making a given statement one is to be properly said to be "giving an analysis" of a *concept*, then (a) both *analysandum* and *analysans* must be *concepts*, and, if the analysis is a *correct* one, must, in some sense, be *the same concept*, and (b) that the *expression* used for the *analysandum* must be a different *expression* from that used for the *analysans*... and a third may be added: namely this: (c) that the *expression* used for the *analysandum* must not only be *different* from that used for the *analysans*, but that they must differ in this way, namely, that the expression used for the *analysans* must *explicitly mention* concepts which are not explicitly mentioned by the expression used for the analysandum.... And that the *method of combination* should be explicitly mentioned by the expression used for the *analysans* is, I think, also a necessary condition for the giving of an analysis.[18]

Now it seems to us that Moore is on the right track in insisting on conditions (a), (b), and (c) — and, of course, in adding as a fourth condition a specification of the method of combination (what we have called "propositional combination"). Where he went wrong, we suggest, was in his attempt to give these conditions a summary description by saying, at the outset, that "one must suppose that both statements [the informative one and the noninformative one] are in some sense *about* the expressions used as well as *about* the concept of being a brother" [our italics]. For once he had put it this way, the puzzle was for him — as indeed it would be for us — insoluble. He had already rejected the view — as we have — that a statement of analysis is about verbal expressions: "both *analysandum* and *analysans* must be concepts or propositions, *not* mere verbal expressions."[19] Little wonder that he could not see clearly in *what* sense the analysans is "about" the verbal expressions used. It is not "about" them in any sense whatever.

16. Arthur Pap, *Semantics and Necessary Truth,* New Haven, Yale University Press, 1958, p. 275.

17. Among the most interesting are those of Alonzo Church, "The Paradox of Analysis", *Journal of Symbolic Logic* (1946) pp. 132–3 (discussed critically by Pap in *Semantics and Necessary Truth,* pp. 276–279); Pap himself, *op. cit.;* and John L. Mackie, *Truth, Probability and Paradox,* Oxford, Clarendon Press, 1973, pp. 1–16.

18. *The Philosophy of G.E. Moore*, p. 666.

19. *Ibid,* p. 664.

The desiderata we have to keep in mind, if a solution is to be found, emerge more clearly when we compare the following three sentences:

(4.11) "The expression 'being a brother' expresses the same concept as (means the same as) the expression 'being a male sibling'."

(4.12) "Being a brother is identical with being a male sibling."

(4.13) "Being a brother is identical with being a brother."

The first thing to note is that these sentences are *not* all "about" the same thing. *(4.11)* is about the verbal expressions "being a brother" and "being a male sibling": it says of them that they express the same concept (or, as it is sometimes put, "mean the same"). *(4.12)* and *(4.13)*, however, are not about verbal expressions but about concepts. On this score, then, *(4.11)* fails to meet Moore's condition (a), whereas both *(4.12)* and *(4.13)* satisfy it. The second thing to note is that these sentences do *not* all express analyses. *(4.11)* fails again because, as Moore points out, it merely tells us that two expressions express the same concept without telling us what that concept is: it "could be completely understood by a person who had not the least idea of what either expression meant".[20] And *(4.13)* fails because it does not meet either of Moore's conditions (b) or (c). Sentence *(4.12)*, it is clear, is the only sentence which satisfies all three conditions, (a), (b), and (c) — and, for that matter, the fourth condition which concerns the method of combination as well. The third point to note is that these sentences do *not* all seem equally informative. Sentence *(4.11)* certainly gives us information, even though it is informative about words not about concepts. Sentence *(4.12) seems* to be equally informative; but the trouble is to see how it can be since, on the one hand, it is not about words but concepts, and on the other hand, it expresses the very same proposition as does *(4.13)* which plainly is not informative about anything at all.

The solution which eluded Moore, though he came very close to finding it, is really quite simple. It lies in recognizing that *a sentence can be informative about something which that sentence is not about* — or more particularly, that a sentence can give information about a verbal expression even when it is not *about* that verbal expression (does not *mention* it) but simply *uses* that verbal expression to say something about a concept. Thus although the sentence

(4.12) "Being a brother is identical with being a male sibling"

is about the concepts of being a brother and being a male sibling (it truly asserts their identity) and not about the verbal expressions "being a brother" and "being a male sibling" (which plainly are *not* identical), nonetheless this sentence *contains* the expressions "being a brother" and "being a male sibling" among its constituents; it *uses* these expressions to say something about the concepts of being a brother and being a male sibling; and, *in using them, it conveys information about the use of these expressions,* viz., the information that they may be used to express one and the same concept. If this is correct, we can then explain why it is that *(4.12)* seems just as informative as *(4.11)* even though it is not about what *(4.11)* is about. *(4.12)* is as informative as *(4.11)* because *(4.12) conveys* (or *shows*) what *(4.11) says*.[21] And we can also explain why it is that *(4.13)* seems totally

20. *The Philosophy of G.E. Moore*, p. 662.

21. The distinction between what sentence expresses (i.e., says) and what it conveys (i.e., shows) should be easily grasped at the intuitive level. It can, however, be explicated in strictly logical terms. To say that *(4.12) shows* that the expression "being a brother" expresses the same concept as the expression "being a male

uninformative even though it expresses exactly the same proposition as does *(4.12)*. For what *(4.13)* shows is *not* the same as what *(4.12)* shows. It conveys only that the expressions "being a brother" and "being a brother" may be used to express one and the same concept — which is not to convey any information at all. In short, *(4.12)* says what *(4.13)* says but does not show what *(4.13)* shows; and at the same time, *(4.12)* shows what *(4.11)* says but does not say what *(4.11)* says.[22] By virtue of what *(4.12)* says, it expresses identity; by virtue of what it shows, it is informative.

The above solution seems to be perfectly general insofar as it applies not only to the paradox of analysis but also to the wider paradoxes of identity. If the number 9 is identical with the number 3^2, how can the equation "$9 = 3^2$" be more informative than the trivial equation "$9 = 9$"? It would be informative, of course, to say that the numeral "9" expresses the very same number as does the expression "3^2". But, the equation "$9 = 3^2$" does not say this: it says something about numbers, not about the numerals or other devices which express them. Our solution applies straightforwardly. The equation "$9 = 3^2$" says just what the equation "$9 = 9$" says but does not show what "$9 = 9$" shows (viz., something uninformative). And at the same time, the equation "$9 = 3^2$" shows what is said when we utter the sentence "The numeral '9' expresses the very same number as does the expression '3^2' ", but it does not say what this sentence says.

A correct statement of analysis (i.e., a sentence expressing a correct analysis), like any correct statement of identity, *can* tell us something which we do not already know: it can tell us that certain expressions express the same concept (number, proposition, or whatnot), not indeed by *saying* anything about them, but by *showing,* by the way we use them in the sentence expressing the analysis, that they express the same concept (number, proposition, or whatnot). The method of analysis has nothing to fear from the so-called "Paradox of Analysis". It is not a genuine paradox at all but a solved puzzle.

4. THE METHOD OF INFERENCE

Inference, it may be recalled, may play a role both in the acquisition of experiential knowledge and in the acquisition of ratiocinative knowledge. Thus, we saw (in chapter 3) that experiential knowledge is knowledge gained either by direct appeal to experience or *by inference therefrom* while ratiocinative knowledge is knowledge gained by appeal to reason, e.g., by analysis of concepts or *by inference therefrom.* In each case, inference serves as a means whereby knowledge which has already been gained may be *expanded.*

But what exactly is inference? Does all inference yield knowledge? How can we be sure that any particular inference will yield knowledge rather than mere true belief, or, worse still, false belief? Only when we have answered these and other related questions will we be able fully to understand the role which inference plays in the science of logic.

---

sibling" is to say that the proposition (1), that a person who in uttering the sentence *(4.12)* expresses the necessary truth that being a brother is identical with being a male sibling, *implies* the proposition (2), that the expressions "being a brother" and "being a male sibling" express the same concept.

22. Students of Frege and the early Wittgenstein will recognize that this solution owes something to Frege's distinction between sense and reference and to Wittgenstein's distinction between what can be said and what can be shown. Nevertheless, Frege's distinction (which was, incidentally, invoked by Church in his attempt to solve the paradox) is not ours. Neither is Wittgenstein's. For Wittgenstein, what is said is always something contingent, what is shown is always something noncontingent, and what is said can never be shown. For us, by way of contrast, one can show something contingent and one can say it too. To repeat: *(4.12)* shows what *(4.11)* says, viz., something contingent about the uses of the expressions "being a brother" and "being a male sibling".

The nature of inference

We would do well, for a start, to get clear as to what we mean by "inference". Even so usually reliable a source as *Webster's New Collegiate Dictionary*[23] can get it wrong. One of the definitions offered there of "inference" is

> the act of passing from one proposition, statement or judgment considered as true to another whose truth is believed to follow from that of the former.

This will not do. In the first place, an act of inference need not be performed on a proposition which is "considered true".[24] It may be performed on a proposition whose truth-value is quite *unknown* to us; as, for example, when we test a hypothesis by drawing inferences from it and then checking to see whether the propositions inferred are true or false. And it may also be performed on a proposition which is considered, or even known, by us to be *false;* as, for example, when — in cross-examination — we demonstrate that a certain piece of testimony which we already believe or know to be false, really is false by showing that certain patently false propositions can validly be inferred from that testimony. The source of *New Collegiate*'s error is not too hard to diagnose. Ordinary usage, in most contexts, sanctions the insertion of the expression "the truth of" before the expression "the proposition that" (and equally before such expressions as "the statement that", "the belief that", etc.). Equally it sanctions the deletion of this expression. Strictly speaking, in most cases the presence of "the truth of" is redundant — except perhaps for emphasis or stylistic flourish — and also quite innocuous.[25] Thus, the two expressions "infer Q from P" and "infer *the truth of* Q from *the truth of* P" are synonymous. This allows us to say, for example, that the inference from the proposition that the moon is made of green cheese to the proposition that the moon is edible *is the very same inference* as the inference from *the truth of* the proposition that the moon is made of green cheese to *the truth of* the proposition that the moon is edible. The latter use of the expression "the truth of" loses its innocuousness only if, as was done in the *New Collegiate*, we not only suppose — as we should — that all cases of drawing an inference from a proposition P are cases of drawing an inference from the truth of a proposition P but take the further step of supposing — as we should not — that all such cases are cases of drawing an inference from a proposition P "considered as true". This further (inferential) step is quite unwarranted. If it is made, it obscures the important fact that inferences can be, and are, validly made from propositions regardless of their truth-values and regardless, too, of the beliefs, if any, that we happen to have about their truth-values. What is true is not that when one infers Q from P one believes P and, on that basis, believes Q, but rather that when one infers Q from P, then *if* one believes that P *then* one believes that Q. The first amendment that is called for, then, is the deletion of the words "considered as true".

The second amendment needed is the deletion of the word "another". Must the proposition inferred be different from the proposition from which it is inferred? Consider the inference from the proposition

23. Springfield, Mass., G. & C. Merriam Co., 1974.

24. *Webster's New Collegiate Dictionary* is not alone in making this first mistake. Philosophers have often made it, too. For instance, Peter Strawson tells us that in inferring or drawing conclusions "you know some facts or truths already, and are concerned to see what further information can be derived from them". (*Introduction to Logical Theory*, London, Methuen, 1952, p. 13.) Similarly, Stephen Barker defines it thus: "*Inference* is the mental act of reaching a conclusion from one's premises, the achievement of coming to believe the conclusion because one comes to see, or think one sees, that it follows logically from premises already accepted as true." (*The Elements of Logic,* New York, McGraw-Hill, 1965, p. 8.)

25. Note that the two sentences, "The book is on the table" and "It is true that there is a book on the table", express logically equivalent propositions. This case is generalizable for all sentences which express propositions.

(4.14) There are at least 10 marbles in the bag,

to the proposition,

(4.15) There are no fewer than 10 marbles in the bag.

Most of us would regard the inference from *(4.14)* to *(4.15)* as a valid one despite the fact that *(4.14)* is *the very same proposition* as *(4.15)*. Yet, if we were to take the *New Collegiate* definition seriously, we should have to deny that this is a valid inference on the grounds that it is not an inference at all. This time the source of error is likely to be found in custom, or metaphor, or both. It is undoubtedly true that the inferences we customarily draw lead us from one proposition, or set of propositions, to another (non-identical) proposition or set of propositions. And it is undoubtedly true that the seemingly apt metaphor of "passing from" — as it is employed in the above definition — encourages the belief that where there is inference there is passage, movement, or at least difference. But that which is customary need not — as the inference from *(4.14)* to *(4.15)* makes clear — be universal. And metaphor, when it is misleading, should be abandoned. In the present instance, neither custom nor metaphor should deter us from allowing the propriety of saying that a proposition may be inferred from itself.

The *New Collegiate* definition of "inference" can withstand the first of our suggested amendments. But with the attempt to make the second its syntax falls apart. We therefore offer the following preliminary definitions of "inference":

> "inference" $=_{\text{df}}$ "an act or a series of acts of reasoning which persons perform when, from the truth of a proposition or set of propositions, P, they conclude the truth of a proposition or set of propositions, Q".

Or again, capitalizing on the redundancy of "the truth of" in both of its occurrences in the above, more simply:

> "inference" $=_{\text{df}}$ "an act or a series of acts of reasoning which persons perform when, from a proposition or set of propositions, P, they conclude a proposition or set of propositions, Q".

Several points about our repaired definition deserve comment.

On the one hand, the repaired definition avoids the errors of the *New Collegiate*'s definition. The question is left open, as it should be, as to whether in any given inferece the proposition P, from which inference is made, is true (let alone considered, believed, or known to be true). And the question is also left open as to whether in any given inference, the proposition Q, which is inferred from some proposition P, is the same as or different from P. (Consistently, throughout this book, we use the letters "P" and "Q" as propositonal *variables* to stand for any propositions whatever, even one and the same proposition.)

On the other hand, our definition preserves the merits of that in the *New Collegiate*. It endorses, indeed further emphasizes, the fact that inference is a human act (or at least is an act of a conscious, reasoning creature). This fact is important insofar as, like any human act, the act of inference is subject to evaluation: in particular — and this is a point which we will develop shortly — it is an act which is sometimes performed well and sometimes performed badly.

In spite of these improvements, our preliminary definition is not wholly satisfactory. It neglects the fact that inference may properly be made from, and to, concepts as well as propositions. For example, we should want to say that we can make an inference from the concept of being a brother to the concept of being a male. To get it quite right, the definition of "inference" should be expanded by adding:

"or when, from a concept C_1 they conclude a concept C_2."

It should be emphasized that it is propositions and concepts from which, and to which, inferences are made. To be sure, it is perfectly proper to speak of inferring that someone is angry from his raised voice, red face, and violent gesticulations. But, on analysis, we should want to say that the inference in such a case still takes propositions or concepts as its point of departure — that, strictly speaking, it is not the raised voice, the red face, and violent gesticulations from which we make the inference but rather from the person's voice being raised, his face being red, his gesticulations being violent, etc. And a similar construction, we suggest, is to be placed upon such talk as that of "inferring from the evidence", "inferring from the silence", and so on. Indeed, our own earlier accounts of experiential and ratiocinative knowledge likewise admit of propositional reconstrual. When we say that experiential knowledge is knowledge gained either by direct experience or by inference therefrom, we mean that it is knowledge gained by experience either by direct experience or from *propositions* known thereby. And similarly, when we say that ratiocinative knowledge is knowledge gained, for example, by analysis of concepts or by inference therefrom, we mean it is knowledge gained, for example, by analysis of concepts or by inference from *propositions* known thereby.

We are all probably aware of the fact — to which our definition of "inference" gives due recognition — that inference is a practical human activity, that some of us are better at it than others, and that even the most skillful inferrers do not always draw their inferences unerringly. In short, we recognize that some inferences are "good" (or, as we shall prefer to say, *valid*) while others are "bad" (or, as we shall prefer to say, *invalid*).

But what is it for an inference to be valid as opposed to invalid? How can we be sure that any given inference is valid rather than invalid? And how, in practice, can we proceed so as to maximize inferential acuity and minimize inferential error?

Valid and invalid propositional inferences

The question as to how valid inferences differ from invalid ones is primarily a logical one. Not surprisingly, its answer may be obtained, ratiocinatively, by reflecting on the very definition of "inference". Since propositional inferences are acts which persons perform when from a proposition or set of propositions P, they conclude a proposition, or set of propositions Q, any inference from P to Q will be a valid one only if Q follows from P, and will be an invalid one only if Q does *not* follow from P (chapter 1, p. 32). Sharpening up this answer in definitional form we have:

> "The immediate inference from P to Q is valid" $=_{df}$ "Q is inferred solely from P and Q follows from P"

and

> "the immediate inference from P to Q is invalid" $=_{df}$ "Q is inferred solely from P and Q does not follow from P".

Some inferences, however, are not immediate: one infers P from Q, not solely from P, but through the mediation of some other proposition or propositions. Thus we have:

> "The mediate inference from P to Q is valid" $=_{df}$ "Q is inferred from P through a *sequence* of immediate inferences each of which is valid"

and

> "the mediate inference from P to Q is invalid" $=_{df}$ "Q is inferred from P through a *sequence* of immediate inferences one or more of which is invalid".

Now it is important to note that there is not just one notion of what it is for Q to follow from P: there is a broad notion and there is a narrow one. In the narrow sense of the word, when we say that Q follows from P we mean that Q *follows necessarily* from P. In the broad sense of the word, when we say that Q follows from P we mean that Q *follows probably* from P. And according to whether we choose to interpret the expression "follows from" in one way or the other, the expressions "valid" and "invalid", as they occur in the above definitions of valid and invalid inference, may themselves be construed either narrowly or broadly.

The inference from P to Q is said to be *deductively valid* when Q follows in the narrow sense, i.e., necessarily, from P. It is in this narrow sense of "follows from" that we were using the expression when, in chapter 1 (p. 31), we said that the relation of following from is the converse of the relation of implication. Thus the inference from P to Q is deductively valid just when P *implies* Q.

The inference from P to Q is said to be *inductively valid* when Q follows in the broad sense, i.e., follows probably, from P (or, conversely, when P *probabilifies* Q).

Hitherto we have not had occasion to use the expressions "follows from" and "valid" in any but their narrow, deductive senses. The broader, inductive senses of these words will be discussed more fully in chapter 6. There we shall see that these broader notions, like their narrower counterparts, can be defined in terms of the notion of *possible worlds* and hence that the inductive relation of *probabilification*, like its deductive cousin, *implication*, is a *modal* notion.[26] For the most part, however, our concern in this book is with deductive logic, with deductive validity, with following necessarily and its converse, implication. Accordingly, unless express notice is given to the contrary, we shall *continue* to use words like "logic", "validity", and "follows from" in their narrow, deductive senses while occasionally adding the qualifier "deductive" by way of reminder.

Now it is all very well to be told what logical properties an inference must have in order for that inference to be valid. It is quite another matter to be able to *ascertain* whether any particular inference has those logical properties. How can this be done?

Determining the validity of inferences: the problem of justification

Many philosophers who would unhesitatingly answer the logical question about the nature of validity as we have express puzzlement — even concern — about how this further epistemic question should be answered.

The problem of justifying any claim that an inference is *inductively* valid is notoriously difficult. Ever since it received its classic formulation in the writings of David Hume,[27] the problem of induction has perplexed many and even haunted some. It comes in many guises. How can we be sure that the so-called 'laws of probability' will not change? Why, if at all, is it reasonable to employ certain inductive rules of inference? And so on. There seems not yet to be any generally accepted solution to this problem (or rather, this nest of problems).[28]

26. Roughly, P *probabilifies* Q if *most* of the possible worlds in which P is true (if any) are worlds in which Q is true. By way of contrast, of course, P *implies* Q if and only if *all* of the possible worlds in which P is true (if any) are also worlds in which Q is true.

27. See David Hume, *A Treatise of Human Nature,* ed. L.A. Selby-Bigge, Oxford, Clarendon Press, 1888, book I, part 3, and *An Enquiry Concerning Human Understanding,* ed. Eric Steinberg, Indianapolis, Hackett Publishing Co., 1977, especially sect. 4. A recent restatement is to be found in Nelson Goodman's *Fact, Fiction, and Forecast,* 3rd ed., Indianapolis, Hackett Publishing Co., 1979, chapters 3 and 4.

28. For a succinct review of the problem and various attempts at its solution, see Max Black's article, "Induction", in the *Encyclopedia of Philosophy,* ed. Paul Edwards, London, Macmillan, 1967, vol. 4, pp. 169–181.

Likewise, the problem of how we can ever justify the claim that an inference is *deductively* valid has seemed to some both a problem and a source of potential philosophical embarrassment. The range of proffered solutions to a large extent parallels that for the problem of induction. Thus, as Nelson Goodman has pointed out in his *Fact, Fiction, and Forecast,*[29] some philosophers have thought that rules of deductive logic can be justified by appeal to higher-order rules, principles or axioms; some have thought that they are grounded in the very nature of the human mind; and some, Goodman among them, have thought that they can be justified pragmatically, as it were, in terms of their efficacy in leading us to conclusions which we find acceptable.

The solution seems, in principle at least, to be fairly straightforward. If, as we have said, an inference from P to Q is deductively valid just when P implies Q, we can *justify* the claim that an inference from P to Q is deductively valid simply by *justifying* the claim that P implies Q. And that is something which, so far as we know, can be done in one way and one way only, viz., a priori, by *analyzing* the conditions for the application of the concepts involved in the propositions concerned. For instance, we can justify the so-called Rule of Simplification, which says that from the assertion of P and Q one may validly infer P, by showing that the corresponding statement of implication, viz., that the proposition that P and Q are true *implies* the proposition that P is true, is analytically true. To be sure, analysis does have its limits. Some noncontingent propositions, including statements of implication, seem to resist all attempts at analytical justification; and even in cases where analytical knowledge can be gained it is by no means always gained easily. To the extent that there are limits to our ability to determine, by analysis, whether or not a proposition P implies a proposition Q, there will be limits also to our ability to determine whether the inference from P to Q is deductively valid. But — and this is the important point — the limits to the latter do not exceed the limits to the former. The problem of justifying inference is not endemic: it does not apply to inferences per se but to those only for which the corresponding statements of implication cannot be justified by analysis.[30] Of course, the question can always be raised as to how recourse to the method of analysis is itself to be justified. But that, it is clear, is another question. In any case, the fact is that, despite any deep doubts that may linger about the ultimate justification of inference, analysis, or whatnot, there is a large measure of agreement over which inferences are deductively valid and which are not; and, when pressed to justify these logical appraisals, persons standardly take recourse to analytical methods.

Let us, then, agree that we can and do often justify claims that our inferences are deductively valid and that we can do this in the last resort, by analyzing the propositions from which and to which our inferences are made to see whether the logical relation of implication really does hold between them.

Even so, a problem remains. Although analysis is the ultimate guarantor of inferential validity, it does not serve as a *practical* guide whereby we can direct our inferential activity. Compare the predicament of a would-be mountain climber who knows what it is to be atop a mountain, and who knows how — if called upon — he would justify his claim to have climbed it, but who still wants to know: How do I get there from here (without a slip)? The problem remaining for a would-be inferrer, after he has been told what it is for an inference to be valid and how — if called upon — he might justify his claim to have performed a valid inference, may be expressed in the same way: How do I get there from here (without a slip)? What is required, in each case, is a set of principles or rules for achieving the desired end. In the case of the would-be mountain climber, the rules are rules of climbing and will be treated under such headings as "Belaying", "Rappelling", and the like.

29. pp. 63–64.

30. Note that it does not follow from anything which we have said that a similar solution can be found for the problem of justifying claims that certain inferences are inductively valid. On this question the controversy is too large for us to pursue.

In the case of the would-be inferrer, the rules are rules of inference and go by such names as "Simplification", "Modus Ponens", and the like.

Rules of inference

The formulation of deductively valid rules of inference has long been a standard occupation of logicians. Needless to say, sound inferential practice did not have to await the formulation of such rules. Well before Aristotle first formulated the rules of syllogistic reasoning, persons constructed valid syllogisms. And even today, countless persons perform countless valid inferences each day in complete ignorance of the rules that contemporary logic affords. The essential point about rules of inference is not that we must *know* them in order that we might perform our inferences validly, but only that when we do perform our inferences they must accord with these rules if they are to be valid. Moreover, if we *do* know them, our inferential performances are — like those of the mountain climber who knows the rules for belaying, rapelling, etc. — likely to be improved. Knowledge of the rules of valid inference helps — as we put it earlier — to maximize inferential acuity and minimize inferential error. Knowledge of rules of valid inference can help us to infer from a given proposition only what follows from that proposition.

Now it is a common feature of all rules — rules of inference included — that they admit of universal application. By this we do not mean that rules necessarily have more than one actual instance of application; we mean only that rules necessarily have more than one *possible* instance of application. It makes no sense to speak of a rule which *can* apply to one case only.

It is, of course, just because rules are universal, and admit of application to indefinitely many instances, that they have the pragmatic value which logically minded persons have always cherished from Aristotle to the present day. In order to ascertain whether any *particular* inference that we want to make or have already made is valid or not, we need not undertake the painstaking and often difficult business of deeply *analyzing* the particular premises and conclusion that feature in that inference, checking to see whether there are any possible worlds in which those premises are true while that conclusion is false, so as to determine whether those particular premises imply that particular conclusion. We need to analyze the propositions only to the extent that we can ascertain whether or not that particular inference is an instance which accords with one of those general rules of inference which logicians have *antecedently certified,* by analysis, as valid.

As examples of some of the most commonly invoked rules of inference we cite a handful from the Logic of Unanalyzed Propositions.

Conjunction:

From the proposition P and the proposition Q, one may validly infer the conjunction of P and Q.

[Example of application: Nixon knew about Watergate. Agnew knew about Watergate. ∴ Nixon knew about Watergate and Agnew knew about Watergate.]

Simplification:

From the conjunction of P and Q, one may validly infer the proposition P.

[Example of application: Mao Tse-Tung was an opponent of Chiang-Kai-Shek and Chou-En-Lai was an opponent of Chiang-Kai-Shek. ∴ Mao Tse-Tung was an opponent of Chiang-Kai-Shek.]

Identity:

From the proposition P, one may validly infer the proposition P.[31]

[Example of application: It is going to rain. ∴ It is going to rain.]

Addition:

From the proposition P, one may validly infer the disjunction of P with Q.

[Example of application: John Doe will have heard the news. ∴ Either John Doe will have heard the news or Sue will have heard the news.]

Transposition (sometimes known as *Contraposition*):

From the proposition that if P then Q, one may validly infer that if it is not the case that Q then it is not the case that P.

[Example of application: If John is married to Sue then John is Sue's husband. ∴ If it is not the case that John is Sue's husband, then it is not the case that John is married to Sue.]

Modus Ponens:

From the proposition that if P then Q and the proposition P, one may validly infer the proposition Q.

[Example of application: If all caged animals are neurotic then Felix is neurotic. All caged animals *are* neurotic. ∴ Felix is neurotic.]

Modus Tollens:

From the proposition that if P then Q and the proposition that Q is false, one may validly infer the proposition that P is false.

[Example of application: If all caged animals are neurotic then Felix is neurotic. Felix is *not* neurotic. . . Not all caged animals are neurotic.]

Hypothetical Syllogism (also known as the *Chain Rule*):

From the proposition that if P then Q and the proposition that if Q then R, one may validly infer the proposition that if P then R.

[Example of application: If it rains the snow will melt. If the snow melts the World Cup slalom will be cancelled. ∴ If it rains the World Cup slalom will be cancelled.]

Disjunctive Syllogism:

From the proposition that either P or Q and the proposition that P is false, one may validly infer Q.

31. The recognition that Identity is a valid rule of inference should dispel any lingering doubts about our criticism of the definition given of "inference" in *Webster's New Collegiate Dictionary*.

[Example of application: Either John Doe will have heard the news or Sue will have heard the news. John Doe will not have heard the news. ∴ Sue will have heard the news.]

Constructive Dilemma:

From the three propositions, that if P then Q, that if R then S, and that either P or R, one may validly infer the proposition that either Q or S.

[Example of application: If Thoeni wins the slalom an Italian will win the World Cup. If Stenmark wins the slalom a Swede will win the World Cup. Either Thoeni or Stenmark will win the slalom. ∴ Either an Italian or a Swede will win the World Cup.]

What kind of rule is a rule of inference?

The class of rules, of whatever kind, can be subdivided into two mutually exclusive and jointly exhaustive subclasses: the class of rules which are propositions and the class of rules which are nonpropositions. The class of nonpropositional rules includes such rules as "Keep off the grass" and "Do not drink and drive." Of such rules, it makes no sense to ask whether they are true or false. The concepts of truth and falsity do not apply to such entities.[32] The class of propositional rules includes rules such as "All residents earning a gross annual income in excess of $1,500 are required to file an income tax return." Of this latter rule, it *is* perfectly proper to ask whether it is true or not.

Are rules of inference propositional or nonpropositional ones? Both what we have already said of inference rules and a perusal of the examples just given should make the answer clear. Inference rules are *propositions,* and it is proper to ask of alleged or proffered inference rules whether they are true or false. Consider, for example, the Rule of Simplification: from P and Q one may validly infer P. Whatever else one might wish to say of this rule, one thing which cannot be gainsaid is that it is *true.*

Although inference rules are undoubtedly *useful,* the justification of a particular inference rule does not lie in citing its utility; it lies *only* in an a priori demonstration of its *truth.*

To say that valid inference rules are *true* is to say something about their logical status. It is not to describe the manner in which they may be *used* in making inferences. This latter matter requires that we look at another way of classifying rules.

The class of all rules, of whatever kind, may be subdivided into a second set of mutually exclusively and jointly exhaustive subclasses: the class of rules which are directives and the class of rules which are nondirectives. Among the directive propositional rules are to be found such rules as "To multiply a number by eleven, first multiply that number by ten and then add the product to the original number." Among the nondirective propositional rules are to be found such rules as "On the anniversary date of the mortgage, the mortgagor may make payments in multiples of $1,000 against the principal balance." Of both of these kinds of propositional rules we can properly inquire as to their truth-values. But there is a difference between them: the former may be viewed as a 'recipe'; the latter not. The former kind of rule tells us explicitly how to proceed in a given circumstance; it outlines an explicit series of steps to be followed. The latter is not a recipe; it merely tells us what *may* be done, i.e., what is permitted, in a certain circumstance.

32. This is not to say, of course, that there might not be propositions which 'correspond' to these rules, such propositions as, e.g., that the city's bylaws forbid one's walking on the grass, or that there is a $1,000 fine for drinking and driving, etc.

Now which of these two kinds of rules are rules of inference? Are they recipes, or do they merely state what it is logically permissible to do?

A brief examination of the foregoing list of inference rules provides an answer. Inference rules are not recipes; each tell us only that some particular inference is permitted; none tells us which particular inference we should make.

Thus it turns out that talk of 'following' an inference rule is entirely inappropriate. One can 'follow' a set of recipes; but inference rules are not recipes.

One cannot 'follow' a set of inference rules in order to make a valid inference. Rather it is that one reasons and makes inferences, and if the inferences are valid, then they are properly said to *accord with* valid rules of inference.

To every valid inference there corresponds a valid rule of inference with which that inference accords. But this is not to say that to every valid inference there corresponds a *known* valid rule of inference. Inferences may be valid without our knowing that they are and without our being able to cite a known valid inference rule with which they accord.

Inference and the expansion of knowledge

That the making of a valid inference may lead us to knowledge of propositions which we did not previously know is fairly obvious. What is not obvious, however, is precisely *why* this is so, under what *conditions* it is so, and what *sort* of knowledge valid inference-making can yield.

The class of all cases in which we make a valid inference from a proposition P to a proposition Q may be divided into two subclasses:

(1) the cases in which we make the inference and *know* that the proposition P is true;

(2) the cases in which we make the inference and do *not* know that the proposition P is true.

Consider, first, the cases in which we validly infer Q from P and know P to be true. Now we have already seen, from our discussion of the nature of knowledge (chapter 3, section 2), that if a proposition P is known to be true then it *is* true. And we have also seen, from our discussion of the nature of valid inference (earlier in this section), that if we validly infer Q from P where P is true, then Q is true. It follows, then, that if we validly infer Q from P where P is known to be true, Q is *true*. But does it also follow, in these circumstances, that Q will be *known* to be true? And if not, what further conditions need to be satisfied in order for Q to be known?

First let us remind ourselves of the four conditions which we found, by analysis, to be separately necessary and jointly sufficient for the application of the concept of knowledge: we can be said to know P if and only if P is *true,* we *believe* that P is true, our belief that P is true is *justified,* and this justified true belief is *indefeasible.* It will help, for illustrative purposes, if we pursue our inquiry in terms of an example.

Let P be the (compound) proposition

(4.16) If it rains the snow will melt, and if the snow melts the World Cup slalom will be cancelled

and let Q be the (compound) proposition

(4.17) If it rains, the World Cup slalom will be cancelled.

Clearly, the immediate inference from *(4.16)* is (deductively) valid. It accords with the rule of inference which we called Hypothetical Syllogism. Suppose that we *know* *(4.16)* is true. Does it follow, when we validly infer *(4.17)* from *(4.16)*, that we *know* *(4.17)* to be true?

Well, in the first place, it is easy to show that *(4.17)* satisfies the first condition for its being known, viz., that *(4.17)* is *true*. This follows from the fact that since, by hypothesis, *(4.16)* is known to be true, *(4.16)* is true, together with the fact that if *(4.16)* is true and *(4.17)* is validly inferred from *(4.16)* — as it is — then *(4.17)* is true. Secondly, *(4.17)* will be *believed* to be true. Two conditions suffice for our believing *(4.17)*: one, that we infer *(4.17)* from a proposition P; two, that we believe P to be true. Both these conditions are satisfied when we infer *(4.17)* from *(4.16)*. As may be recalled, although we disagreed with the view that inference is always from believed propositions to believed propositions, we subscribe to the view that when inference *is* made from a proposition which is believed to be true, then this inferred proposition is also believed to be true. Thirdly, our true belief in *(4.17)* will be *justified*. It is justified insofar as it accords with the antecedently certified rule of Hypothetical Syllogism. Logical appeal of this kind counts as a paradigm of justification.

Up to this point we have shown that the first three of the four necessary conditions for knowing Q are satisfied. But is the indefeasibility condition also satisfied? One thing that is clear is this: that since *(4.17)* follows from *(4.16)* and *(4.17)*is true, there cannot be any true proposition R, belief in which would equally justify our concluding that *(4.17)* is false. For in order that we should be equally justified in concluding that *(4.17)* is false on the basis of a belief in R, R would not only have to be true but be such that the falsity of *(4.17)* follows from R. But this is impossible. There cannot be any two *true* propositions, from one of which it follows that *(4.17)* is true while from the other of which it follows that *(4.17)* is false. However, this does not entitle us to conclude that the indefeasibility condition is satisfied. In order that it should be satisfied there must not be any true proposition R, belief in which would even *undermine* the belief that *(4.17)* is true. Suppose, then, that R is the proposition that we came to make the valid inference from *(4.16)* to *(4.17)*, not on the grounds that it accords with the Rule of Hypothetical Syllogism, but on the mistaken grounds of a belief that it accords with the Rule of Modus Ponens. And suppose, further, that R is true. Then, if we were to *believe* that R is true, i.e., were to believe that our inference from *(4.16)* to *(4.17)* has been made on mistaken grounds, our belief in the truth of *(4.17)* *would* be undermined. Hence, the defeasibility condition would not, in the circumstances envisaged, be automatically satisfied.

Plainly, then, we are not entitled to conclude that in the case where we make a valid inference from P to Q and know P to be true we will also know Q to be true. But we *are* entitled to conclude that in such a case we will know Q to be true *provided that* the defeasibility condition happens also to be satisfied.

The conclusion we have just established about all cases of kind (1) is, of course, perfectly general. It matters not *how* we obtained knowledge of P in the first instance, whether experientially or ratiocinatively. It is worth noting, however, that from our earlier definitions of "experiential knowledge" and "ratiocinative knowledge" it follows that if we know Q by validly inferring it from P where P is known experientially then Q will also be known experientially; whereas if we know Q by validly inferring it from P where P is known ratiocinatively, then Q will also be known ratiocinatively.[33]

Before turning to the examination of cases of kind (2) — cases in which we validly infer Q from P and do *not* know that P is true — in order to determine what sort of knowledge, if any, such

33. This is not to suggest, however, that Q (like P) might not be known in both ways. Our argument, of chapter 3, pp. 168–170, showed that some noncontingent propositions can be known in both ways.

inference may yield, it is worth observing that this kind of case covers a variety of possibilities. In saying that we do not know that P is true we are allowing the possibility of our merely believing, but not knowing, that P is true; the possibility of our not having the faintest idea whether P is true or not; the possibility of our believing, but not knowing, that P is false; the possibility of our knowing that P is false; and even the possibility of our knowing that P is necessarily false. The question before us is simply whether in any of these cases, our validly inferring Q from P can yield knowledge, and, if so, knowledge of what.

On first hearing the question we may be inclined to answer: No — that mere inference by itself cannot yield knowledge, i.e., that inference cannot yield knowledge unless it is inference from a proposition, or propositions, which are already known. But this, we shall now try to show, would be a mistake.

Our answer will be, to the contrary, that in cases of kind (2), valid inference may indeed yield knowledge: knowledge not, perhaps, of the truth of Q; but knowledge, rather, of the conditional proposition that if P is true then Q is true.[34]

Once more, for illustrative purposes, we shall argue the case in terms of a particular (arbitrarily chosen) example. This time, for the sake of variety, let P be the (compound) proposition

> *(4.18)* If all caged animals are neurotic then Felix is neurotic; and all caged animals are neurotic

and let Q be the (simple) proposition

> *(4.19)* Felix is neurotic.

Clearly, the immediate inference from *(4.18)* to *(4.19)* is (deductively) valid. It accords with the rule of inference which we called Modus Ponens. Suppose that the truth of *(4.18)* is not known (perhaps because we do not know the truth of one of its conjuncts, viz., that all caged animals are neurotic). Under what conditions, when we validly infer *(4.19)* from *(4.18)*, does it follow that we *know* the truth of the proposition that if *(4.18)* is true then *(4.19)* is true?

Our discussion goes along the same broad lines as that given for cases of kind (1). In the first place, the proposition that if *(4.18)* is true then *(4.19)* is true is a *true* proposition. This, as already noted, follows trivially from the fact that *(4.19)* may validly be inferred from *(4.18)*. Secondly, the proposition that if *(4.18)* is true then *(4.19)* is true, will be *believed* to be true. This follows from the fact that when we infer *(4.19)* from *(4.18)* we *conclude* that *(4.19)* is true if *(4.18)* is true, which is — *inter alia* — *to believe* that *(4.19)* is true if *(4.18)* is true. Thirdly, our true belief that if *(4.18)* is true then *(4.19)* is true will be *justified.* It is justified by appeal to a valid rule, viz., Modus Ponens, with which the inference accords. But will the inference be indefeasibly justified? It may be, but it need not. Suppose someone who makes this inference by appeal to Modus Ponens does so fortuitously — e.g., not out of an understanding that this rule sanctions this particular inference, but purely out of a habit to invoke this particular rule, a habit which is exercised more often incorrectly than correctly. In such circumstances the indefeasibility condition would not be satisfied and we would want to say that one's justified true belief in the conditional proposition that if *(4.18)* then *(4.19)* was *not* knowledge. Nonetheless, in cases of kind (2), the indefeasibility condition may, equally well, *be* satisfied. One's inference may well proceed, not from mere habit of

34. Note that we are not saying that where Q is validly inferred from P then, independently of whether or not we know P to be true, it follows that if P is true then Q is true. To say that is to say something true but — in the light of the intimate connection between the notions of validity and of implication — it is also rather trivial. Our thesis is the stronger one that (subject to one proviso) where Q is validly inferred from P then, independently of whether or not we know P to be true, it follows that we *know* that if P is true then Q is true.

thought, but from genuine understanding — and if so, then it is possible that one should gain knowledge.

More generally, we may conclude, from our examination of this case, that our validly making an immediate inference to a proposition Q, from a proposition P, in circumstances in which it is possible to cite an antecedently certified rule of inference, is by itself sufficient — provided the defeasibility condition is satisfied — to give us knowledge of the truth of the conditional proposition that if P is true then Q is true. This conclusion, it should be noted, also holds for inferences of kind (1).

The *sort* of knowledge obtained in cases of kind (2) is importantly different from that obtained in cases of kind (1). Whereas in cases of kind (1) the knowledge obtained of Q may be *either* experiential (where P is known experientally) *or* ratiocinative (where P is known ratiocinatively), and the knowledge of the conditional proposition that if P is true then Q is true is likewise ratiocinative, the knowledge obtained in cases of kind (2) is *always ratiocinative,* i.e., is always obtained solely by the exercise of one's own powers of reason, and hence is *always* a priori.[35]

But whichever kind of valid inference is involved, one thing is clear: the making of a valid inference may lead us to knowledge of the truth of propositions which we did not previously know to be true; and it can do this with respect to any field of human knowledge whatever, or with respect to propositions belonging to any subject matter whatever.[36]

EXERCISES

For each of the following say with which rule of inference it accords.

1. *Eva is a bank director. Joseph is a lawyer.* ∴ *Eva is a bank director and Joseph is a lawyer.*

2. *Harry is a judge.* ∴ *Harry is a judge or Harry is a former district attorney.*

3. *If Paul is older than Lorna then Lorna isn't ahead of him in school. Paul is older than Lorna.* ∴ *Lorna isn't ahead of Paul in school.*

35. It may be wondered, at this point, why we have assigned inference a subsidiary role in our account of how ratiocinative (and hence a priori) knowledge can be gained. Recall that we defined ratiocinative knowledge as knowledge obtainable "by appeal to reason, e.g., by analysis of concepts or by inference therefrom". Why, it may be asked, the word "therefrom"? Have we not just shown that the making of a deductively valid inference can, all by itself, yield knowledge of certain propositions? Why, then, do we not count inference and analysis as *two,* independent, means to ratiocinative knowledge? The answer lies in the fact that we wish to accord analysis a certain *epistemic primacy* which — on our view — inference does not possess. By this we mean that, if our earlier arguments are sound, it is sound analysis which is the ultimate guarantor of the validity of inferences, not the validity of inferences which is the ultimate guarantor of sound analysis. To be sure, we must, in the course of analyzing a concept or proposition, make inferences; and these inferences, if the analysis is to be sound, must be valid. To that extent, analysis and inference go hand-in-hand and will stand or fall together (as sources of ratiocinative knowledge). But it seems to us that there is a very good sense in which, when it comes to matters of justification, the method of analysis serves to justify the method of inference, not the other way around. If we are wrong about this, our definition of ratiocinative knowledge can easily be repaired by the simple expedient of dropping the word "therefrom".

36. Although the arguments have been conducted by means of examples of immediate inferences, our conclusions can be generalized to the case of mediate inferences as well.

4. *If Martin is older than Jennifer, then Jennifer is older than Jonathan. If Jennifer is older than Jonathan, then Jonathan is older than Diane. ∴ If Martin is older than Jennifer, then Jonathan is older than Diane.*

5. *If it is necessary that all aunts are females then it is possible that all aunts are females. It is necessary that all aunts are females. ∴ It is possible that all aunts are females.*

5. INFERENCE WITHIN THE SCIENCE OF LOGIC

Among the various uses of rules of inference are their uses in *deductive systems.* By a "deductive system" we mean a body of proposition-expressing sentences or formulae which is systematized by means of certified deductively valid rules of inference in such a way as to display logical interconnections between its various items. Although deductive systems can be constructed within fields other than logic — notably mathematics (Euclid's *Elements* provides the first known example of a deductive system) and parts of physics — it is within the science of logic (sometimes called "the science of inference") that they have their purest form. Most commonly, deductive systems are either *axiomatic systems* or *natural deduction systems.* Our aim, in what follows, will be first to show how these two kinds of deductive systems may be employed within *logic* and then to show how their employment therein can lead to the systematic expansion of *logical knowledge.*

Axiomatic systems for truth-functional propositional logic[37] and predicate logic were first constructed by Frege in his *Begriffsschrift* (1879) and developed more fully by Whitehead and Russell in their *Principia Mathematica* (3 vols., 1910–1913). The construction of axiomatic systems for modal propositional logic was first essayed by C. I. Lewis in his *Survey of Symbolic Logic* (1918) and then developed more fully in his *Symbolic Logic* (coauthored with C. H. Langford, 1932). It is to one of the systems offered in the latter book, the system known as S5, that we turn for illustrative purposes. We have chosen the axiomatization of a *modal* logic for reasons of principle, having to do with the philosophical standpoint of this book. Since logic is the study of modal properties and relations, only a *modal* logic can be expected to do philosophical justice to the subject of logic. We have chosen S5, from among the several systems which Lewis developed (and the numerous systems which others have developed), because it seems to us, as to many others, that it is the 'strongest' philosophically defensible system of propositional modal logic, insofar as every other modal logic which subsumes S5 contains theses which seem to us philosophically indefensible as explications of implication, consistency, possibility, necessity, etc.[38]

Inference within axiomatic systems: the example of S5

Perhaps the first thing to note is that although S5, like any other axiomatic system, could in principle be constructed solely with the use of expressions in a *natural language* such as English, no axiomatic system ever is thus constructed. Rather, in S5, as in all other axiomatic systems, recourse is taken to a *symbolic language*[39] in terms of which the truths belonging to the subject matter

37. What it is for a logic to be truth-functional will be explained in detail in chapter 5. Roughly, a propositional logic is truth-functional if the truth-values of its compound propositions are determined by, or a function of, the truth-values of its unanalyzed constituent propositions.

38. For a lengthy discussion of its defensibility, see W. and M. Kneale, *The Development of Logic,* Oxford, Clarendon Press, 1962, pp. 548–568.

39. There are several reasons why recourse is invariably taken to symbolism. One is precisely that which

concerned may be expressed, while the use of natural language is restricted to our descriptions of how the truths of that subject matter may initially be expressed and subsequently generated. We call the symbolic language in terms of which the truths of the system are expressed the *object-language* of the system, and the language, in terms of which we talk about expressions in the symbolic object-language, the *meta-language* of the system.

We commence our sketch of S5 by using English as our meta-language in order to describe the *axiomatic basis* of that system. An axiomatic basis for S5[40] comprises:

(a) A list of the *symbolic vocabulary* to be employed in the object-language. Some of the symbols are taken as undefined or "primitive" while others are defined. Thus we may list as our primitive symbols:

"P", "Q", "R", etc. (propositional symbols)

"$\sim$", "$\cdot$" (symbols for the concepts of negation and conjunction, respectively)

"$\Diamond$" (symbol for the concept of logical possibility)

and go on to define further symbols thus:

$$\text{"}(P \vee Q)\text{"} =_{df} \text{"}\sim(\sim P \cdot \sim Q)\text{"}$$

$$\text{"}(P \supset Q)\text{"} =_{df} \text{"}\sim(P \cdot \sim Q)\text{"}$$

$$\text{"}(P \equiv Q)\text{"} =_{df} \text{"}((P \supset Q) \cdot (Q \supset P))\text{"}$$

$$\text{"}\Box P\text{"} =_{df} \text{"}\sim \Diamond \sim P\text{"}$$

$$\text{"}(P \rightarrow Q)\text{"} =_{df} \text{"}\sim \Diamond(P \cdot \sim Q)\text{"}$$

$$\text{"}(P \leftrightarrow Q)\text{"} =_{df} \text{"}((P \rightarrow Q) \cdot (Q \rightarrow P))\text{"}$$ [41]

---

has led us, increasingly throughout this book, to use symbols. What one wants to say may thereby be expressed more succinctly and unambiguously. Another is that artificially introduced symbols are usually free from the disease of association-of-ideas — the disease which usually infects our uses of expressions in natural languages and bedevils the inferences which we try to make in terms of them.

40. We here take some liberties with Lewis' account by simplifying and employing — in some cases — different symbols and terminology.

41. The symbols "$\cdot$" (for conjunction), "$\vee$" (for disjunction), "$\supset$" (for material conditionality), and "$\equiv$" (for material biconditionality) deserve comment. Together with the symbol "$\sim$" (for negation) they compromise the standard repertoire of truth-functional symbols. The concepts of negation, conjunction, and disjunction have already been defined; but a brief reminder is in order. The *negation* of a given proposition is true in all those possible worlds (if any) in which that proposition is false, and false in all those possible worlds (if any) in which that proposition is true. The relation of *conjunction* hold between two propositions in all those possible worlds (if any) in which both are true. The relation of *disjunction* holds between two propositions in all those possible worlds (if any) in which at least one of them is true. The concepts of material conditionality and material biconditionality are readily definable along the lines of Lewis' definitions. Thus we shall say that the relation of *material conditionality* holds between P and Q in all those possible worlds (if any) in which it is

(b) A set of *formation rules.* These are, in effect, rules of grammar for the symbolic language being constructed — rules, that is, which determine which formulae constructed out of the symbolic vocabulary are to count as grammatical, i.e., *well-formed formulae* (wffs), and which are to count as ungrammatical, i.e., *not well-formed.* The formation rules of the system S5 are:

R1: Any propositional symbol standing alone is a wff.

R2: If α is a wff, so is $\sim\alpha$.[42]

R3: If α is a wff, so is $\Diamond\alpha$.

R4: If α and β are wffs, so is $(\alpha \cdot \beta)$.

(We do not need to give formation-rules for wffs involving the defined symbols since these are effectively provided by the definitions for the introduction of these symbols.) By successive applications of these rules we can generate all and only the well-formed formulae of our symbolic language.

(c) A selected set of wffs, known as *axioms.* A great deal of care usually goes into selecting, from among the infinitely many wffs which the formation rules allow us to construct, the relative handful which are to be accorded the privileged status of axioms. For purposes of constructing S5, Lewis selected A1 to A6 plus A10 out of the following list:

A1: $(P \cdot Q) \rightarrow (Q \cdot P)$

A2: $(P \cdot Q) \rightarrow P$

A3: $P \rightarrow (P \cdot P)$

A4: $((P \cdot Q) \cdot R) \rightarrow (P \cdot (Q \cdot R))$

A5: $((P \rightarrow Q) \cdot (Q \rightarrow R)) \rightarrow (P \rightarrow R)$

A6: $P \rightarrow \Diamond P$

A7: $\Diamond(P \cdot Q) \rightarrow \Diamond P$

A8: $(P \rightarrow Q) \rightarrow (\Diamond P \rightarrow \Diamond Q)$

---

not the case that P is true and Q false, while the relation of *material biconditionality* holds between P and Q in all those possible worlds (if any) in which both are true or both are false. Subject to the qualifications discussed in chapter 5, section 2, "$\sim$P" may be read as "not P", "P · Q" as "both P and Q", "P∨Q" "either P or Q", "P ⊃ Q" as "if P then Q", and "P≡Q" as "P if and only if Q".

42. Note that the Greek letters "α", "β", etc. do not belong to the object-language but to the meta-language. They are used to supplement the meta-language, English, and are known as *meta-logical variables.* "α", "β", etc. stand indiscriminately for any wffs whatever.

A9: $\Diamond\Diamond P \rightarrow \Diamond P$

A10: $\Diamond P \rightarrow \Box\Diamond P$

(A1 to A6, he showed, suffice to construct a very weak system which he called S1. Progressively stronger systems are constructible by adding further axioms to those for S1. Thus S2 = A1 to A6 + A7; S3 = A1 to A6 + A8; S4 = A1 to A6 + A9; and S5 = A1 to A6 + A10. Although A7, A8, and A9 do not feature as *axioms* in S5 they are provable as theorems therein.)

(d) A set of rules of inference, known as *transformation rules*. The transformation rules of a system enable us to transform the axiom-wffs into new wffs, and these in turn into still further wffs. Any wff obtained in this way is known as a *theorem* of the system. (Together with the axioms, the theorems make up the *theses* of the system.) It is usual to be as parsimonious as possible in selecting one's transformation rules. Thus, for the purposes of deriving the theorems of S5 we can get along nicely with:

TR1: [Conjunction] if α is a thesis and β is a thesis, then $(\alpha \cdot \beta)$ is a thesis.

TR2: [Modus Ponens] If α is a thesis and $(\alpha \rightarrow \beta)$ is a thesis, then β is a thesis.

TR3: [Uniform Substitution] If α is a thesis and β is the result of substituting some wff for a propositional symbol uniformly throughout α, then β is a thesis.

TR4: [Substitution of Equivalents] If α is a thesis in which β occurs, and $(\beta \leftrightarrow \gamma)$ is a thesis, and one substitutes γ for some occurrence of β in α, then the result of that substitution is a thesis.

Of these, TR1 and TR2 are already familiar (though we earlier stated them slightly differently). TR3 and TR4 are unfamiliar, but may easily be understood in terms of examples of their application. Thus TR3 licences us to make such inferences as those from

A6: $P \rightarrow \Diamond P$

to each of the following:

T1: $Q \rightarrow \Diamond Q$

T2: $(P \cdot Q) \rightarrow \Diamond(P \cdot Q)$

and so on. In effect, TR3 reflects the fact that since a propositional variable is simply a wff which is arbitrarily chosen to stand for any proposition whatever, we can substitute for it any other arbitrarily chosen proposition-expressing wff whatever, provided that we do so consistently, i.e., uniformly. TR3, unlike TR4, does not require of any two wffs, one of which is to be substituted for the other, that they be equivalent. On the other hand, TR4, unlike TR3, does not require of a wff which is to be substituted for another, that it be substituted uniformly. Thus, for instance, supposing that we have already established, as a thesis, the equivalence

$$\text{T3:} \quad P \leftrightarrow \sim\sim P,$$

TR4 licences us to make the inferences from T3 and

$$\text{A6:} \quad P \rightarrow \Diamond P$$

to

$$\text{T4:} \quad \sim\sim P \rightarrow \Diamond \sim\sim P$$

(in which the substitution of "$\sim\sim P$" for "P" *is* carried out uniformly), or to

$$\text{T5:} \quad \sim\sim P \rightarrow \Diamond P$$

(in which the substitution is *not* carried out uniformly).

From the above axiomatic basis for S5 the rest of the system may be generated by repeated applications of one or more of the rules of inference listed in (d) to one or more of the axioms listed in (c) or to one or more of the theorems previously so generated.

The general concept of *proof* may be given the following rigorous definition for the special case of S5: a finite sequence of formulas (A through T), each of which either (i) is an axiom of S5, or (ii) is a theorem derived from one or more previous members of the sequence in accord with a stated transformation rule of S5, is said to be a *proof of* T *in S5*.[43]

There are infinitely many theorems derivable in S5. We list just a few of special interest.

$$\text{T6:} \quad \Box P \rightarrow P$$

$$\text{T7:} \quad (P \rightarrow Q) \rightarrow (\Box P \rightarrow \Box Q)$$

$$\text{T8:} \quad \Diamond(P \cdot Q) \rightarrow (\Diamond P \cdot \Diamond Q)$$

$$\text{T9:} \quad (\Box P \vee \Box Q) \rightarrow \Box(P \vee Q)$$

$$\text{T10:} \quad \sim\Diamond P \rightarrow (P \rightarrow Q)$$

$$\text{T11:} \quad \Box P \rightarrow (Q \rightarrow P)$$

$$\text{T12:} \quad (\Box P \cdot \Box Q) \rightarrow (P \leftrightarrow Q)$$

$$\text{T13:} \quad \Box P \leftrightarrow \Box\Box P$$

$$\text{T14:} \quad \Diamond P \leftrightarrow \Box\Diamond P$$

$$\text{T15:} \quad \Box P \leftrightarrow \Diamond\Box P$$

Questions involving the understanding of T6 to T12 will be posed as an exercise. T13 and T14 are

43. For illustrations of S5 proofs see p. 221ff.

singled out because of their role as so-called "reduction principles". Their philosophical significance will be discussed later.

It is worth noting that corresponding to each of the above theses of S5 there is what is called a "derived inference rule of the system". Roughly, a derived rule (or "derived transformation rule", as it is sometimes called) is a rule of inference which does not occur in the original set but which corresponds to an axiom or to an already established theorem and hence can be derived from the original set. Since to every theorem there corresponds a derivable rule of inference and there is an unlimited number of theorems, we may conclude that there is an unlimited number of derived rules of inference to be obtained if one wishes. Of course, nobody ever so wishes. A system encumbered by as many derived rules of inference as theorems would be ridiculously redundant. In practice, one makes use of a relatively small number of derived rules.

Inference within natural deduction systems

Turning now to deductive systems of the second kind, viz., natural deduction systems, it will suffice for our purposes if we concentrate mainly on the description of how such systems operate.

Natural deduction systems were first constructed independently by Gerhard Gentzen and Stanisław Jaśkowski in 1934. The employment of such systems in various branches of formal logic — propositional and predicate logics, modal, and nonmodal — is now well established. Indeed, for a variety of reasons, they are now regarded by many logicians with more favor than axiomatic systems. The troublesome task of selecting just the 'right' axioms as the starting point for our deductions is circumvented. Both the kinds of inference rules employed,[44] and the manner of their employment are much more natural (as the term "natural deduction" is intended to betoken) than in the case of axiomatic systems. And they do not tempt us so strongly to suppose that the subject matter of logic is on a par with that of, e.g., Euclidean geometry or classical mechanics, by virtue of being likewise axiomatizable — a supposition which fails to recognize that logic has a special status as the science which provides rules of inference for the systematic investigation of these other sciences whereas they provide none for it.

Although we have no need here to give an illustration of how a full system of natural deduction is set up it will help if we illustrate how the rules of such a system operate by applying such rules to an example expressed first in natural language and then subsequently in the symbolism of truth-functional propositional logic.

Consider the argument of someone who asserts all three of the following propositions:

(4.20) If Stenmark did not win the slalom then he did not race

(4.21) But either he raced or the snow conditions must have been dangerous and he withdrew

(4.22) Stenmark did not win

and goes on to conclude

(4.23) The snow conditions must have been dangerous.

44. The rules usually are not selected so austerely and usually are not as seemingly artificial as, e.g., the rule of Uniform Substitution.

Most of us can 'see' that the argument is valid. But how do we *know*? How could we *demonstrate* that the conclusion follows from the premises? An appeal to axiomatics is obviously not going to be much help. No premise is an instance of any axiom, nor is the conclusion an instance of any theorem of any axiomatic system of logic. To be sure, the conditional proposition that if the premises are true then the conclusion is true will, if the argument is valid, be an instance of a theorem of an axiomatic system of logic. But to show that it is would be a lengthy and difficult matter and a quite unnatural thing to do in the circumstances. How much more natural it would be for us to reason as follows:

From the premises *(4.20)* and *(4.22)* it follows that

(4.24) Stenmark did not race.

But from *(4.24)* and the other premise, *(4.21)*, it follows that

(4.25) The snow conditions must have been dangerous and Stenmark withdrew

from which it obviously follows that

(4.23) The snow conditions must have been dangerous.

What we are doing in this piece of natural deduction is implicitly invoking various familiar rules of inference in order to bridge the gap from premises to conclusion by constructing a *series of steps*, each of which follows from one or more of the preceding steps or premises, and the last step of which is the conclusion. We could, if we wished, make it quite explicit as to which rules of inference we are invoking. If so, we could point out that the rule which gets us from the conjunction of *(4.20)* and *(4.22)* to *(4.24)* is the rule of *Modus Ponens;* that the rule which gets us from the conjunction of *(4.24)* and *(4.21)* to *(4.25)* is the rule of *Disjunctive Syllogism;* and that the rule which gets us from *(4.25)* to the conclusion *(4.23)* is that of *Simplification*. In so doing, we would be demonstrating the validity of the argument and the necessary truth of the corresponding conditional, by showing, in step-by-step fashion, that the conclusion follows from (may be derived from, may validly be inferred from) the premises.

The example just given does not strictly belong to any natural deduction system since it is not expressed in a purely symbolic language and, as we have pointed out, no deductive systems — axiomatic or otherwise — are ever expressed in anything but symbolic notation. On the other hand, it is clear that the example just given *could* have been expressed in some set of symbols for propositional logic and that, when so expressed, the derivations involved *could* have been made in a quite mechanical way. If it had been so expressed the argument would look as follows:

$$\left.\begin{array}{ll} (4.20a) & \sim P \supset \sim Q \\ (4.21a) & Q \vee (R \cdot S) \\ (4.22a) & \sim P \end{array}\right\} \text{Premises}$$

$$\therefore \quad (4.23a) \quad R \qquad\qquad \text{Conclusion}$$

Our *proof* that *(4.23a)* follows from the conjunction of the premises would then be set out thus:

	(4.24a)	$\sim Q$	From *(4.20a)* and *(4.22a)* by Modus Ponens
	(4.25a)	$R \cdot S$	From *(4.21a)* and *(4.24a)* by Disjunctive Syllogism
$\therefore$	*(4.23a)*	R	From *(4.25a)* by Simplification,

where the explicit justification for each step in the proof is set out on the right-hand side by citing both the rule of inference which warrants the step and the previous steps (either premises or subsequent steps or a combination of the two) from which the inference is made.

The crucial difference between natural deduction systems and axiomatic systems can now be brought out. In an axiomatic system we begin with theses (the axioms) and we end with theses (the theorems), and all the intermediate steps are theses (viz., theorems). Every single step in the proof of a thesis must itself be a thesis. In a system of natural deduction this need not be the case. We *may* begin with a thesis (i.e., a thesis of an axiomatic system); but we *need not.* And, even when we begin with a thesis and end with a thesis, some of the intermediate steps need not be theses.

It is noteworthy that, in our descriptions of these two kinds of deductive system we have not needed to invoke any talk of the *truth* of the theses they generate, let alone talk of our *knowledge* of the truth of such theses. This reflects the important fact, noted earlier (p. 205, fn. 39), that both kinds of deductive system are invariably constructed with the help of a symbolic notation within which derivations of theses may be constructed without any potentially misleading distractions of the kind that so often plague our inferences when we have a particular interpretation in mind. Although deductive systems are usually *constructed* with some interpretation of the symbols in mind, they may be *considered* independently of any such interpretation. When so considered, a deductive system is said to be an *uninterpreted system.*

Now it is clear that within an uninterpreted system it is *pointless* to ask of any given thesis whether that thesis expresses a truth, let alone whether what it expresses is known to be true. Such questions simply do not arise. Nevertheless they can be *made* to arise — can be made pointful — provided that certain sorts of interpretations are assigned to the symbols, i.e., just as soon as the system is made an *interpreted system.* Strictly speaking, from a formal point of view an uninterpreted axiom is nothing other than a string of marks on a paper or a string of sounds. These marks or noises can be made to express propositions by our interpreting them, i.e., by our specifying for each of their constituent symbols what meaningful interpretation it is to have. Consider for example, the string of symbols "(P $ Q) # (P ! Q)". We stipulate that "$", "#", and "!" are dyadic operators, that "(" and ")" are disambiguating punctuation, and that "P" and "Q" are variables. Beyond this, nothing more is specified. What does this axiom express? Literally, nothing. But it can be *made* to express an indefinitely large range of propositions. We will illustrate just two.

First, the string of symbols may be given a fairly obvious interpretation which would have it express a truth of physics, viz., that the value of the resultant force brought about by two independent forces, P and Q, acting in the same direction on a point [i.e., "(P $ Q)"] is numerically equal to [i.e., "#"] the arithmetic sum of the two individual forces [i.e., "(P ! Q)"].

Alternatively, the string may be given quite a different interpretation, this time, however, yielding a false proposition of physiological audiometrics, viz., that the perceived pitch of the complex sound consisting of two notes sounding in immediate succession [i.e., "(P $ Q)"] is indistinguishable from [i.e., "#"] the perceived pitch of those same two notes sounding in unison [i.e., "(P ! Q)"].

Thus it is that, for any theses of suitably interpreted systems, we can ask the questions which we have hitherto so studiously shunned. We can ask: Do the theses (or sentences which are their substitution-instances[45]) express true propositions? And further: Can we *know* that the

45. The parenthetical qualification is needed in order to cater for deductive systems whose theses are not themselves proposition-expressing sentences but rather are *sentence-forms* all of whose *substitution-instances* are

theses (or sentences which are their substitution-instances) express true propositions?

These questions can be asked of a thesis of an interpreted deductive system no matter what interpretation is given — whether the interpretation given is the *intended interpretation*, i.e., the interpretation which it was intended that the symbolism should bear when the system was constructed, or some *non-intended interpretation*, i.e., an interpretation which it was not intended that the symbolism should bear but which it can be given. The symbolism of truth-functional propositional logic was first constructed with a logical interpretation in mind: an interpretation according to which the letters "P", "Q", "R", etc. were to be interpreted as *propositional* variables, the symbols "$\sim$", "$\cdot$", and "$\vee$", etc., as expressing the *truth-functional* concepts of negation, conjunction, and disjunction, etc., and so on. It so happens, however, that the symbolism may also be given a different interpretation — an interpretation according to which it does not present a systematization of propositional logic but a systematization of the theory of electrical switching circuits. And other unintended interpretations are also possible.

In what follows we will set aside these unintended, nonlogical, interpretations and concentrate solely on the intended, logical, interpretations which systems such as those for propositional logic and predicate logic are standardly given. We will concentrate, that is, solely on questions about the truth, and our knowledge of the truth, of the logical propositions which, on the intended interpretation, the theses of various deductive systems express.

Consider the case of axiomatic systems. A wff is not usually designated as a thesis unless it has at least a prima facie case for being *considered* true on the intended interpretation. We know that *if* the axioms of a system S are true, i.e., true on the intended interpretation (this is a qualification which we hereinafter take for granted when speaking of the truth of wffs), and the rules of inference are valid, i.e., truth-preserving in all possible worlds, then the theorems will be true. Moreover, we know that if the axioms of S are *necessarily* true, not just contingently so, and the rules of inference are valid, then the theorems will be necessarily true. (This can easily be verified by inspection of the fifteen worlds-diagrams which make it clear that from a necessary truth only necessary truths follow [See chapter 1, p. 51].) But how do we *know* whether the axioms of S are true? This plainly cannot be settled within S itself (even when S is interpreted). After all, among the considerations which dictate our initial choice of the axioms is the requirement that the axioms be *independent* of one another in the sense that, although the theorems are derivable from them in accordance with the rules of inference of the system[46] none of them is itself derivable in this way from the *other*

---

proposition-expressing sentences. The distinction between sentence-forms and the sentences which are their instances is drawn with some care in chapter 5, section 6. Strictly speaking, it makes no sense to ask questions about the truth-value or modal status of the propositions expressed by theses which have the status of sentence-forms rather than sentences. For there are no such propositions. On the other hand, we *can* ask about the truth-value and modal states of the propositions expressed by the sentences which are *substitution-instances* of such theses. A sentence-form all of whose instances express necessarily true propositions will be said to be *valid;* one all of whose instances express necessarily false propositions will be said to be *contravalid;* and one which is neither valid nor contravalid will be said to be *indeterminate.* (See chapter 5, section 7.) Although the distinction will be observed scrupulously throughout chapters 5 and 6, it will render our present discussion simpler if we leave it to be *understood* here that talk of the truth-value or modal status of a thesis is subject to the parenthetical qualification.

46. Note the *special* sense in which it is required that the axioms be "independent". This is a different sense of "independence" from that involved when we said, in chapter 1, that two propositions are *logically independent* if and only if from the truth or falsity of one we cannot validly infer the truth or falsity of the other. Within an axiomatic system two axioms are *inferentially independent* if and only if from the truth or falsity of one we cannot validly infer, *by means of the transformation rules of that system,* the truth or falsity

axioms. If, then, the axioms of S are independent, it is not possible to establish their truth by derivation within S. It follows that the only way in which the truth or necessary truth of the axioms of S can be established is *extrasystematically*. And that, it is clear, is where recourse must be taken to what we have called "the method of analysis".

In the case of truth-functional propositional logic, this is standardly done by giving the sort of truth-condition analysis which is to be found in so-called *truth-tables*.[47] In the case of modal propositional logic, this is standardly done by constructing *semantic tableaux;* or, alternatively, it can be done more intuitively by appeal to *worlds-diagrams* in the manner already sketched in chapter 1 and subsequently developed more fully in chapter 6. Truth-tables, on the one hand, and semantic tableaux and worlds-diagrams, on the other hand, provide *decision procedures* for deciding on the truth-value (and, indeed modal status) of theses within truth-functional propositional logic and modal propositional logic, respectively. Between them, these analytical methods suffice to determine, for any thesis of propositional logic standardly interpreted, whether or not that thesis expresses something true or false, together with the modal status of what that thesis expresses.

In the case of predicate logic, i.e., logic of unanalyzed concepts, the scene is somewhat different. Not only is there no decision procedure for predicate logic (truth-functional or otherwise) *as a whole;* it is provable that there cannot be one.[48] Certain substantial fragments of predicate logic, however, do submit to appraisal by these, or similar, analytical methods.[49] In short, for all those cases in which it is possible to determine the truth-value or modal status of theses of axiomatic systems for propositional and predicate logics, it is possible to do so by recourse to what we have broadly described as the methods of analysis.

How about the theses of natural deduction systems for these logics? These generate no special problem. A natural deduction system for a given branch of logic does not contain different theses from those within an axiomatic system for that branch, but uses rules of inference to organize, or

---

of the other. Theses of an interpreted system may be inferentially independent even when they are not logically independent.

This fact is particularly pertinent to the controversy which surrounds the so-called paradoxes of implication: that a necessarily false proposition implies any proposition, that a necessarily true proposition is implied by any proposition, and that any two necessarily true propositions are equivalent. These theses, which we will discuss at greater length in section 6, pp. 224–230, have sometimes been construed as asserting respectively that any proposition may be demonstrated to follow from a necessarily false proposition, that any necessarily true proposition may be demonstrated to follow from any proposition whatever, and so on. Consequently, it has often been claimed, these theses, if they were true, would make the matter of the demonstration of noncontingent truths in logic and mathematics a trivial matter. For instance, it would mean that in order to prove the necessarily true proposition that the square root of two is not the quotient of any two whole numbers, we need only derive it from the contingent proposition that it is raining by appeal to the fact that the former is a necessarily true proposition and hence is implied by any proposition whatever. But this worry about the potential for trivializing mathematics and logic is unwarranted. For to say that a necessarily true proposition is implied by any other proposition is not to say or to imply that a necessarily true proposition may be validly inferred *by means of the transformation rules of a certain system* from any proposition whatever. Within any non-trivial, epistemically productive, logical system such pairs of propositions as that the square root of two is not the quotient of any two integers and that it is raining will be inferentially independent with respect to the rules of that system even though they are not logically independent.

47. Recall that a truth-condition analysis is what we prefer to call a "possible-worlds analysis". Truth-tables are simply one form of possible worlds analysis. We introduce them in chapter 5.

48. See G.J. Massey, *Understanding Symbolic Logic,* New York, Harper & Row, 1970, pp. 338–9, for a discussion of what he calls "Church's thesis theorem".

49. For an excellent introduction to the use of such methods for two important fragments of predicate logic, see Hughes and Londey, *The Elements of Formal Logic,* London, Methuen, 1965.

systematize, the theses in a different way. So whatever analytical method suffices to give us knowledge of the truth-value or modal status of theses within an axiomatic system will suffice also to give us knowledge of the truth-value or modal status of theses in the corresponding natural deduction system. They are the very same theses.

Our questions about the truth, the modal status, and our knowledge of the truth and modal status, of the propositions which, on standard interpretations, the theses of various deductive systems of logic express, can now be answered straightforwardly. In the first place, to the extent that, by the employment of analytical methods such as those described, we are able to know of a proposition P (expressed by a thesis of an interpreted deductive system) that it is true, we are also able to know of any proposition Q (expressed by a thesis of an interpreted deductive system) which we validly infer from P, that it also is true. Secondly, to the extent that, by the employment of analytical methods, we are able to know of a proposition P (expressed by a thesis) that it is *necessarily true,* we are also able to know of any proposition Q (expressed by a thesis) which we validly infer from P, that it also is *necessarily true.*

The conclusions just reached can, of course, be generalized. They apply not only to propositions expressed by theses within interpreted deductive systems but to propositions expressed in any way whatever. The making of valid inferences within deductive systems is one way of expanding our knowledge of the subject matter of logic. But it is not the only way. Valid inference made outside the compass of deductive systems also leads to the expansion of logical knowledge.

EXERCISES

1. *By using the method of counterexamples, as outlined in the section on* Possible Worlds Parables *in chapter 2, try to show why the converses of T6, T7, T8, and T9 do not hold.*
2. *Reread chapter 1, pp. 50–53. Which worlds-diagrams are illustrations of T11? Which are illustrations of T12?*
3. *Find two different interpretations of "(P $ Q) # (P ! Q)" which yield truths, and two different interpretations which yield falsehoods.*

* * * * *

The theoretical warrant of the method of direct proof

The distinction drawn earlier between mediate and immediate inference (p. 195) is paralleled by a distinction between mediate and immediate proofs. A mediate proof will have a sequence of steps between the premises and the conclusion; an immediate one will not. Each of the preceding proofs has been a mediate one; later will we cite several examples of immediate (one-step) proofs.

All immediate proofs, and some mediate ones, are direct proofs; that is, are proofs in which every step is derived solely from the premises or by a sequence of steps from the premises. By way of contrast, some mediate proofs are indirect (otherwise known as conditional proofs); that is, are proofs in which additional assumptions, not included within the original premise-set, are introduced (see. for example, step (3) in the proof in footnote 63 on p. 227). In what follows we will be concerned solely with direct proofs.

What logical principles justify the construction of mediate direct proofs?[50]

50. We shall not concern ourselves with the corresponding question for indirect proofs.

It may seem obvious that since the rules of inference employed in constructing each of these intermediate steps are valid, the final step in the construction of the proof, viz., the conclusion, must follow from the premises. But can we prove this? In order to do so we need to invoke two metalogical principles, the Augmentation Principle and the Collapse Principle.

It might be thought that the only metalogical principle involved is that of the Transitivity of Implication, viz., that if P implies Q and if Q implies R, then P implies R. It may seem, that is, that if the premise-set of an argument implies some proposition, and if that proposition in turn implies another, then the premise-set implies the latter proposition, and so on. But this principle does not suffice. For the intermediate steps in a mediate proof are often inferred, not from an immediately preceding step, but from more remote earlier steps; moreover intermediate steps are often inferred not from single antecedent steps but from two or more such steps or premises (see, for example, step *(4.25a)* in the proof on p. 212).

A mediate direct proof consists of a set of premises, A_1 through A_m (whose conjunction we shall call "A"), a number of intermediate steps, B_1 through B_n, and a conclusion C. Schematically, a mediate direct proof looks like this:

$$\left.\begin{array}{l} A_1 \\ \cdot \\ \cdot \\ A_m \end{array}\right\} \text{Premises}$$

$$\left.\begin{array}{l} B_1 \\ \cdot \\ \cdot \\ B_n \end{array}\right\} \text{Intermediate steps}$$

$$\left.\begin{array}{l} C \end{array}\right\} \text{Conclusion}$$

Each step, beginning with B_1 and proceeding through and including the last, the conclusion C, is inferred from some one or more premises or antecedent steps.

The two meta-logical principles involved are:

THE AUGMENTATION PRINCIPLE:

If P implies Q, then the conjunction of P with any other proposition, R, also implies Q.

Symbolically we would have: $(P \rightarrow Q) \rightarrow [(P \cdot R) \rightarrow Q]$

THE COLLAPSE PRINCIPLE:

If P implies Q, then the conjunction of P and Q, viz., P · Q, is logically equivalent to P.

Symbolically we would have: $(P \rightarrow Q) \rightarrow [(P \cdot Q) \leftrightarrow P]$

(The proof of these principles is left as an exercise in chapter 6, section 9.)

Let us now see how these two principles can be invoked to solve our current problem. Consider the first of the intermediate steps, i.e., B_1, in the schema for mediate direct proofs. Since it is the first step, no intermediate step precedes it and it must be inferred from some one or more propositions among the premise-set A. Provided that it is inferred in accord with a valid inference rule, then by the very definition of "valid" we are assured that B_1 logically follows from those premises which are cited in its derivation. Now by the Augmentation Principle we know that if B_1 logically follows from *some* of the premises, then it follows from them *all* (that is, from them all taken together). And by the Collapse Principle we know that if B_1 follows from the premise-set A, then the set of propositions consisting of all the premises along with B_1 is logically equivalent to the original set of premises. In effect, then, we can regard B_1 as just another premise, and the proof can now be regarded as looking like this:

$$\left.\begin{array}{l} A_1 \\ \cdot \\ \cdot \\ A_m \\ B_1 \end{array}\right\} \text{Enlarged Premise-set}$$

$$\left.\begin{array}{l} B_2 \\ \cdot \\ \cdot \\ B_n \end{array}\right\} \text{Intermediate steps}$$

$$C \quad \} \quad \text{Conclusion}$$

We now proceed to repeat the same sort of reasoning in the case of the next intermediate step, B_2. In so doing we show that regarding B_2 as just another premise will leave the validity of the argument quite intact: the set of propositions consisting of A_1 through B_2 is logically equivalent to the set consisting of A_1 through A_m. We continue in this stepwise fashion until we have shown that the entire set of propositions A_1 through B_n is logically equivalent to the original premise-set A. Finally we are in a position to infer C. C, like any other step in the proof, will be inferred from some one or more antecedent propositions in the proof. Again we invoke the Augmentation Principle. Insofar as C logically follows from some propositions in the set A_1 through B_n, it also logically follows from the entire set, A_1 through B_n. But we have already shown that the set A_1 through B_n is logically equivalent to the original set, A_1 through A_m. Thus, in inferring C in this fashion, we have shown that C logically follows from the original premise-set.[51]

In sum, by constructing a proof in which a proposition is validly inferred from a set of propositions which are themselves validly inferred from a premise-set, we show that that proposition is implied by that premise-set, and that the corresponding argument (that is, the proof without its intermediate steps) is *deductively valid.*

51. Whatever logically follows from one of two logically equivalent proposition-sets logically follows from the other also. In symbols: $[(P \leftrightarrow Q) \cdot (P \rightarrow R)] \rightarrow (Q \rightarrow R)$.

EXERCISE

Verify the principle of the Transitivity of Implication by the following procedure:

Select from the fifteen worlds-diagrams (figure (1.i)*) all those in which P implies Q. From each of these, construct additional diagrams in which Q implies R (for example, diagram 9 will give rise to three additional diagrams, i.e., those in which Q is equivalent to R, in which Q implies the contingent proposition R, and in which R is necessarily true). Check to see that in every case where P implies Q and Q implies R, P implies R.*

6. A PHILOSOPHICAL PERSPECTIVE ON LOGIC AS A WHOLE

Our discussion of the fruits of analysis led us to adopt a threefold division of logic into

(1) the Logic of Unanalyzed Propositions (Propositional Logic, as we are now calling it);

(2) the Logic of Unanalyzed Concepts (Predicate Logic);

and

(3) the Logic of Analyzed Concepts (Concept Logic).

Of these, (1) and (2) are standardly recognized and well developed— in no small measure because they lend themselves to systematic exploration by means of deductive systems. By way of contrast, (3) is certainly not well developed and is only rarely accorded recognition as a proper part of logic. One of our aims, in this section, is to give a philosophical defence of its inclusion within the science of logic. Our other main aim is also philosophical. We wish to argue for the centrality within logic as a whole (and a fortiori within each of its three main parts) of the so-called "branch" called Modal Logic, and for the centrality, within that "branch", of the truths expressed by the theses of Lewis' system S5. We concentrate on the role which modality plays in Propositional Logic and make only a few remarks about its role within Predicate Logic or Concept Logic (parts of logic which largely fall outside the compass of this book).

The indispensability of modal concepts within propositional logics

A broad view of propositional logic must allow that propositional logic is a genus within which may be found several species and many subspecies. Two of the main species, *truth-functional* logic and *modal* logic, have attracted principal attention. But there are others, chief among which are *epistemic* logic (dealing with relations between such epistemic concepts as those of knowledge and belief), *deontic* logic (dealing with relations between such ethical concepts as those of obligation and permission), and *tense* logics (dealing with relations between concepts such as those of past, present, and future). And within certain of these it must further be allowed that insofar as different axiomatizations of the same sets of theses, or even axiomatizations of different sets of theses, have been constructed, there are — so to speak — many subspecies.

Now it has long been customary, within introductory presentations of propositional logic, to accord truth-functional logic pride of place either by neglecting these other logics altogether or by representing each of these as a mere "accretion" upon the "central core" of truth-functional logic. There have been several reasons for this, some of them good, some of them not so good. Among the good reasons we must cite the facts that the truth-functional concepts of negation, conjunction, disjunction, etc., lend themselves to simple analysis in terms of the conditions for their application; that these concepts play a more *obvious* role in ordinary argumentation and inference than do modal, epistemic, deontic, and temporal concepts; that a propositional logic which avoids non-truth-functional notions lends itself to the construction of extremely *simple* decision procedures; and that a certain pedagogical elegance is achieved if we first present some axiomatic basis for truth-functional logic and then *add* to it the special axioms and rules which are needed in order to handle these non-truth-functional concepts. The not-so-good reasons include the all-too-common belief that the truth-functional concepts which some uses of the words "not", "and", "or", "if . . . then", and "if and only if" express[52] are the only strictly logical concepts, whereas non-truth-functional concepts such as those standardly expressed by phrases such as "it is necessary that", "it is known that", and "it is obligatory that" are all tainted with extralogical impurities; that logic has no need of non-truth-functional-concepts; and that the non-truth-functional concepts — especially the modal ones — are somehow philosophically suspect. But whatever the reasons, this mode of presentation has encouraged, if not generated, the view that modal logic is just one of many accretions on the central, pure, truth-functional core of logic, and that modal logic merely examines the relationships between the modal concepts of necessity, possibility, impossibility, etc. in the same sort of way as, e.g., deontic logic examines the relationships between the deontic concepts of being obligatory, being permitted, being forbidden, etc.

On our view all this is topsy-turvy. Given that logic is concerned — as, since its founding days, it has universally been agreed to be — with formulating principles of *valid* inference and determining which propositions *imply* which, and given that the concepts of validity and implication are themselves modal concepts, it is modal logic rather than truth-functional logic which deserves to be seen as central to the science of logic itself. We do not deny for a moment that logicians can and usually do pursue the task of determining which principles or rules of inference are valid, and which theses of logic may be derived in accordance with these rules, without giving any thought to modal logic or giving explicit recognition to modal concepts in their symbolism. We deny only that they can give a *philosophically* satisfactory account of the notions of validity and derivation without appeal to those modal concepts which it is the province of modal logic to investigate. From a philosophical point of view, we submit, it is much sounder to view modal logic as the indispensible core of logic, to view truth-functional logic as one of its fragments, and to view "other" logics — epistemic, deontic, temporal, and the like — as accretions either upon modal logic (a fairly standard view, as it happens) or upon its truth-functional component.[53]

Now, for anyone who shares this perspective on the matter, the question immediately arises as to *which* modal logic gives the most philosophically adequate account of the set of modal relations of which the relation of implication is a member. For the trouble is that we are seemingly faced with an embarrassment of riches. Even if we restrict ourselves to the "classical" presentation of modal logic by C. I. Lewis we find no fewer than eight distinct systems or logics: five in the series S1 to S5,

52. It is universally agreed that only some uses of these words can be construed truth-functionally. We spend a good bit of time in chapter 5 saying which uses they are.

53. Our view of the matter is shared, it seems, by W. and M. Kneale. In the section of their book *The Development of Logic* which they devote to modal logic, they try to give a "proper appreciation of the status which modal logic has among deductive theories" (p. 557) by presenting logic in a new way. Essentially, that new way (which they admit is "less easy to understand at first encounter" [p. 558]) involves presenting logic as a set of second-order propositions concerning the relation of implication. Compare our account of the subject matter of logic as second-order propositions which ascribe modal properties and relations to other propositions (chapter 3, pp. 129, 175).

and three in the distinct series S6 to S8. Subsequently, other logicians have constructed many more, some lying as it were 'between' members of the Lewis series, others lying 'outside' those series. There are now so many modal logics that the author of a recent book reports, in his Introduction, that he will discuss "literally hundreds" of them.[54] Fortunately, not many of these demand scrutiny for our purposes. Many have been constructed expressly as instruments for the analysis of specific sorts of verbal contexts in which modal expressions are used, and others investigate the consequences of incorporating special assumptions along with more familiar modal laws. Only a relative handful compete for attention as systematic explanations of the concepts of logical implication, of logical necessity, logical possibility, and the like. And, of these, the best candidate, in the view of many philosophers, is Lewis' system S5.[55]

Probably the most serious objections to the view that S5 gives a philosophically defensible account of the most fundamental concepts of logic are those which stem from the presence within it of the formulae which we listed earlier as

$$\text{T10:}\quad \sim\Diamond P \rightarrow (P \rightarrow Q)$$

$$\text{T11:}\quad \Box P \rightarrow (Q \rightarrow P)$$

$$\text{T12:}\quad (\Box P \cdot \Box Q) \rightarrow (P \leftrightarrow Q)$$

$$\text{T13:}\quad \Box P \leftrightarrow \Box\Box P$$

$$\text{T14:}\quad \Diamond P \leftrightarrow \Box\Diamond P$$

T10, T11, and T12 are commonly referred to as "paradoxes of (strict) implication". They are not unique to S5 but are to be found in a number of weaker systems as well, i.e., systems whose axiomatic bases suffice for the proof of proper subsets of the theorems in S5. Indeed, they are found in systems as weak as S2. T13 and T14 are known as the Weak and Strong Reduction Principles, respectively. We will discuss the Reduction Principles first, and then come back to the so-called paradoxes.

Problems about the reduction principles

Even before we consider the problems that are supposed to arise about T13 and T14, it is possible to explain why the former is said to be "weak" and the latter "strong", and to explain, further, why they are jointly called "Reduction Principles". Significantly enough, we can do this without so much as considering for a moment what these wffs might mean, or, what interpretation they might be given.

In saying that T14 is stronger than T13 we mean not only that S5, within which T14 is provable, is stronger than S4, within which T13 is provable, but also that once we have proved T14, it is a fairly straightforward matter to prove T13. We *show* this by constructing a series of three proofs.

First we prove that the Strong Reduction Principle, T14, is a theorem of S5. The distinctive

54. D. Paul Snyder, *Modal Logic and its Applications,* New York, Van Nostrand Reinhold, 1971.

55. See, for instance, W. and M. Kneale, *The Development of Logic,* pp. 559–566. They claim that S5, when generalized in a system which takes implication as fundamental, suffices "for the reconstruction of the whole of logic as that is commonly understood" (p. 563). See also Jaakko Hintikka's conclusion: "The system S5, then, seems to be the best formalisation of our logic of *logical* necessity and *logical* possibility." ["The Modes of Modality" in *Acta Philosophica Fennica,* vol. 16 (1963).]

axiom of S5, viz., $\Diamond P \rightarrow \Box \Diamond P$, (A10), uses the symbol "$\rightarrow$". By way of contrast, T14, viz., $\Diamond P \leftrightarrow \Box \Diamond P$, uses the symbol "$\leftrightarrow$". How do we get from the former to the latter? It will help simplify our proof of T14 if, instead of using Lewis' axiom A6, viz., $P \rightarrow \Diamond P$, we use another formula, viz., $\Box P \rightarrow P$, which suffices for the generation of precisely the same set of theses as A6. Let us call this formula A6*.[56] Our proof of T14 employs the axiomatic basis of S5, viz., the rules TR1, TR2, TR3, and TR4 together with the axioms A1 to A6 + A10 (or the equivalent set A1 to A6* + A10). We obtain T14 from this axiomatic basis by means of a proof set out thus:

(1)	$\Box P \rightarrow P$	[A6*]
(2)	$\Box \Diamond P \rightarrow \Diamond P$	[(1) × TR3 ($\Diamond P/P$)]
(3)	$\Diamond P \rightarrow \Box \Diamond P$	[A10]
(4)	$(\Box \Diamond P \rightarrow \Diamond P) \cdot (\Diamond P \rightarrow \Box \Diamond P)$	[(2), (3) × TR1]
(5)	$\Diamond P \leftrightarrow \Box \Diamond P$	[(4) × Def. $\leftrightarrow$]

[A brief explanation is in order. On each line of the proof we write a numbered wff which is either an axiom — e.g., wffs (1) and (3) — or a theorem — e.g., wffs (2), (4), and (5). To the right of the wff, in square brackets, we write the *justification* for writing that wff. In the case of axioms, merely citing them as such suffices. In the case of theorems, however, we justify our writing them down by citing a transformation rule or a definition which entitles us to derive them from previously numbered wffs, i.e., from wffs which are axioms or wffs which are previously derived theorems. The abbreviation "[(1) × TR3 ($\Diamond P/P$)]" written after wff (2), for instance, tells us that we obtain (2) from (1) by substituting $\Diamond P$ for P in accordance with the Rule of Uniform Substitution.] The wff which appears as a theorem on the last line of the proof is, of course, the Strong Reduction Principle. Q.E.D.

Next, we prove that the distinctive axiom (A9) of S4, viz., $\Diamond \Diamond P \rightarrow \Diamond P$, is provable as a theorem in S5. Once more it will help us simplify our proof if, instead of using Lewis' formulation of A9, we use another formula, viz., $\Box P \rightarrow \Box \Box P$, which is interchangeable with it, and call it A9*. The proof of A9* (= A9) within S5 then runs thus:

(1)	$P \rightarrow \Diamond P$	[A6]
(2)	$\Box P \rightarrow \Diamond \Box P$	[(1) × TR3 ($\Box P/P$)]
(3)	$\Diamond P \leftrightarrow \Box \Diamond P$	[T14]
(4)	$\Diamond \Box P \leftrightarrow \Box \Diamond \Box P$	[(3) × TR3 ($\Box P/P$)]
(5)	$\Box P \rightarrow \Box \Diamond \Box P$	[(2), (4) × TR4 ($\Box \Diamond \Box P/ \Diamond \Box P$)]
(6)	$\Box P \leftrightarrow \Diamond \Box P$	[T15]
(7)	$\Box P \rightarrow \Box \Box P$	[(5), (6) × TR4 ($\Box P/ \Diamond \Box P$)]

56. It may be of interest to note that Gödel's axiomatization of S5 employs a weakened version of A6* rather than A6. See chapter 6, p. 356.

[The abbreviations of the justifications given for (5) and (7), both of which cite TR4, call for comment. On expansion, the justification given for (5) tells us that since, in (4), $\Diamond\Box P$ and $\Box\Diamond\Box P$ have been proved equivalent, we can obtain (5) from (2) by substituting $\Box\Diamond\Box P$ for $\Diamond\Box P$ where the latter occurs in (2), in accordance with TR4, i.e., the Rule for Substitution of Equivalents. Similarly, the justification given for (7) tells us that since, in (6), $\Box P$ and $\Diamond\Box P$ have been proved equivalent, we can obtain (7) from (5) by substituting $\Box P$ for $\Diamond\Box P$ where the latter occurs in (5), in accordance with TR4.] The wff which appears as a theorem on the last line of this proof is, as already noted, interchangeable with the distinctive axiom of S4. Moreover, the proof is an S5 proof since it utilizes the Strong Reduction Principle, T14, in the proof of which we previously utilized the distinctive axiom of S5, viz., A10. Thus we can conclude, at this point, that S5 contains S4.

Finally, we can show that the Weak Reduction Principle, T13, is provable in S4. Its proof utilizes the distinctive axiom of S4, viz., A9* (= A9), as the following demonstrates:

(1)	$\Box P \rightarrow P$	[A6*]
(2)	$\Box\Box P \rightarrow \Box P$	[(1) × TR3 ($\Box P/P$)]
(3)	$\Box P \rightarrow \Box\Box P$	[A9*]
(4)	$(\Box\Box P \rightarrow \Box P) \cdot (\Box P \rightarrow \Box\Box P)$	[(2), (3) × TR1]
(5)	$\Box P \leftrightarrow \Box\Box P$	[(4) × Def.$\leftrightarrow$]

A little reflection on these proofs shows not only that S5 'contains' S4, since the axiomatic basis of S4 is provable therein (see the second proof), but also that once we have proved T14 within S5 (see the first proof), we can easily prove T13. Thus our second proof used T14, at line (3), together with A6, to prove $\Box P \rightarrow \Box\Box P$, at line (7). And our third proof used this result, together with A6* (= A6), to prove T13. It remains only to add the well-known fact that T14 is *not* provable in S4, but *is* in S5[57], in order to conclude that whereas T13 is derivable from T14, T14 is not derivable from T13. Ipso facto, T14 is the stronger and T13 the weaker of the two Principles.

But why are either T13 or T14 called "Reduction Principles"? Once more we can answer the question without recourse to matters of interpretation. T13 says that $\Box P$ is provably equivalent to $\Box\Box P$. Now according to TR4 (the Rule for Substitution of Equivalents), if two wffs are provably equivalent then one may be substituted for the other wherever it occurs. It follows that in any wff in which there is a double occurrence (an iteration) of the symbol $\Box$ — or, for that matter, in any wff which is equivalent to one in which there is a double occurrence of $\Box$ — we can always delete the left-hand occurrence of $\Box$ and so *reduce* the number of its occurrences. When we do this systematically in the way that T13 indicates, we are left in S4 with just twelve distinct, i.e., non-equivalent, irreducible modalities; viz.,

$\Box P$,	$\Box\Diamond P$,	$\Box\Diamond\Box P$
$\sim\Box P$,	$\sim\Box\Diamond P$,	$\sim\Box\Diamond\Box P$
$\Diamond P$,	$\Diamond\Box P$,	$\Diamond\Box\Diamond P$
$\sim\Diamond P$,	$\sim\Diamond\Box P$,	$\sim\Diamond\Box\Diamond P$

57. This can easily be *seen* although its proof is a difficult matter. After all, if $\Diamond P \leftrightarrow \Box\Diamond P$ were provable in S4 then $\Diamond P \rightarrow \Box\Diamond P$ would be also. But if the latter were provable in S4, there would be no difference between S5 and S4 since $\Diamond P \rightarrow \Box\Diamond P$, as we have seen, is the distinctive axiom of S5.

T14 effects a further reduction so that we are left, in S5, with only the first on each of the above four rows, viz.,

$$\Box P, \qquad \sim\Box P, \qquad \Diamond P, \qquad \sim\Diamond P.$$

T13 and T14 are called "Reduction Principles", then, because T13 reduces all wffs containing combinations of the symbols $\Box$ and $\Diamond$ to wffs containing at the most *three* such symbols, while T14 reduces them to wffs containing at most *one* such symbol.[58] In each case, longer strings are permitted but are dispensable.

Now it is one thing to show, as we have, that T14 (and hence also T13) is provable in S5. It is quite another thing to show that these Reduction Principles, on interpretation, are philosophically defensible. That is our next task.

The interpretation which we have been giving the symbolism of S5 should by now be sufficiently familiar. Nevertheless, it bears summarizing. The letters "P", "Q", "R", etc. are taken to stand for propositions; the symbols "$\sim$" and "$\cdot$" are taken to stand for the truth-functional concepts of negation and conjunction, respectively; "$\Diamond$" and "$\Box$" are taken to stand for the modal concepts of logical possibility (truth in at least one possible world) and logical necessity (truth in all possible worlds), respectively; and "$\rightarrow$" and "$\leftrightarrow$" are taken to stand for the modal relations of implication and equivalence, respectively. The question before us, then, is whether, *on this interpretation* of the symbols, T13 and T14 are true.

Consider T13 first. T13 asserts an equivalence or two-way relation of implication, viz., (a) that $\Box P$ implies $\Box\Box P$, and (b) that $\Box\Box P$ implies $\Box P$. There can hardly be any doubt about the truth of (b), for (b) is simply a substitution-instance of the obvious truth that if P is necessarily true then it follows that P is true (expressed in symbols as $\Box P \rightarrow P$).[59] The only question that can seriously be raised is about (a). But (a)'s truth, on reflection, is similarly obvious. For (a) simply tells us that if P is true in all possible worlds, then the proposition *that* it is true in all possible worlds will itself be true in all possible worlds.

T14 yields to the same sort of analysis. It, too asserts a two-way relation of implication, viz., (a) that $\Diamond P$ implies $\Box\Diamond P$, and (b) that $\Box\Diamond P$ implies $\Diamond P$. Once more (b) must be accepted on pain of denying that necessary truths are true. And (a) simply tells us that if P is true in at least one possible world then the proposition *that* it is true in at least one possible world will be true in all possible worlds. If the obviousness of (a) seems elusive, the following argument may help. To deny the truth of (a) would be to assert that P may have the property of being true in at least one possible world even though the assertion that it has this property is not true in every possible world, i.e., is false in at least one possible world. But this seems obviously false if we think in terms of examples. Let P be the proposition

(4.26) We are at the beginning of a new Ice Age.

Then, whether or not P is true, it is at least *possibly* true. So the antecedent of (a) is true for this substitution-instance of P. Might there, then, be a possible world in which it is false that *(4.26)* is logically possible, i.e., a possible world in which *(4.26)* is logically impossible? Hardly. It now seems obvious that a world in which *(4.26)* is logically impossible, i.e., self-contradictory, is not a

58. On a different account of what is to count as a modality, some logicians count both P and $\sim$P as additional modalities, and hence S4 and S5 have respectively, 14 and 6 irreducible modalities.

59. Obviously, if we allow the truth of $\Box P \rightarrow P$ then we must allow the truth of $\Box\Box P \rightarrow \Box P$ since the latter follows from the former in accordance with the Rule of Uniform Substitution (substituting $\Box P$ for P).

possible world but an impossible one. It follows that one cannot consistently assert the antecedent of (a) and deny its consequent. In brief, on our interpretation, (a) can be seen — on reflection and analysis — to be just as incontrovertibly true as (b).

Why, it may then be asked, have T13 and T14 seemed to some philosophers to be obviously false or even meaningless? An explanation is called for.

The charge of meaninglessness, though it is often enough heard, does not deserve to be taken very seriously. To the rhetorical question, "What could it possibly *mean* to say such things as that it is true in all possible worlds that it is true in all possible worlds that P?", one can only reply that one ought not to let one's mind be boggled by complex strings of words (or symbols) but ought, rather, to set oneself the task of thinking through — as we did above — what they *do* mean.

By way of contrast, the objection which has it that T13 and T14 are obviously false is of considerable philosophical interest since it usually stems from a subtle assimilation of the notions of necessary truth and provability. Suppose, for instance, that instead of interpreting $\Box P$ in T14 to mean that P is necessarily true, we interpret it to mean that P is provable by appeal to a specified set of inference rules, e.g., within a deductive system. Then, since $\Diamond P$ is definitionally equivalent to $\sim \Box \sim P$, we should have to interpret $\Diamond P$ as meaning that P cannot be disproved by appeal to that set of rules. And then the formula $\Diamond P \leftrightarrow \Box \Diamond P$ will be read as asserting that if (and only if) P cannot be disproved by invoking certain rules, then the fact that it cannot be disproved by invoking these rules can itself be *proved* by invoking them. But this, for most cases, turns out to be false. On this interpretation, then, the Strong Reduction Principle of S5, and hence S5 itself, turns out not to be philosophically defensible. The answer that is called for in this case is that the concepts of necessary truth and provability are not at all the same; that the former is a purely logical concept while the latter is an epistemic one; and that unless one takes pains to keep the two distinct — as we have done at length in chapter 3 and again in this chapter — wholesale philosophical confusion is likely to ensure. The objection itself is a case in point. More particularly, we should point out that although being proved to be necessarily true is a sufficient condition of being necessarily true, it is not a necessary condition since there may be — and probably are, if our arguments in chapter 3 are sound — many necessary truths which are neither proved nor provable.

The objections to the Reduction Principles of S5 center around the interpretation, or misinterpretation, of the symbols "$\Box$" and "$\Diamond$". The objections to the so-called paradoxes, T10, T11, and T12, center around the interpretation, or misinterpretation, of the symbol "$\rightarrow$". It is to these that we now turn.

Problems about the paradoxes

Recall, for a start, the interpretations which we have given of T10 through T12.

$$\text{T10:} \quad \sim \Diamond P \rightarrow (P \rightarrow Q)$$

is to mean that if P is necessarily false or self-contradictory then it follows that P implies any proposition Q whatever; i.e., that a necessarily false proposition implies any and every proposition.

$$\text{T11:} \quad \Box P \rightarrow (Q \rightarrow P)$$

is to mean that if P is necessarily true then it follows that P is implied by any proposition Q whatever; i.e., that a necessarily true proposition is implied by any and every proposition. And

$$\text{T12:} \quad (\Box P \cdot \Box Q) \rightarrow (P \leftrightarrow Q)$$

is to mean that if P and Q are both necessarily true then they mutually imply one another; i.e., that necessary truths are equivalent to one another. Since the usual criticisms concentrate on T10 and T11, we will deal first with them and then come back to T12.

One of the commonest complaints about T10 and T11, and hence also about all the Lewis systems, including S5, which contain them, is that they reflect a "highly artificial", "specialized", "purely formal" concept of implication, a concept which bears only a remote resemblance to our 'ordinary' notion of implication. It may help a little, then, if we briefly set the Lewis systems into historical perspective and then say why he thought his systems captured the essential features of the ordinary concept. Right at the outset, let it be admitted that Lewis' own terminology may have contributed, at least in part, to the supposition that there is a 'gap' between ordinary implication and the notion that features in his systems. He called the relation which we have symbolized by "$\rightarrow$", and which he symbolized by the so-called fish-hook "$-\!3$", the relation of *strict* implication. In so naming it, he may well have made it sound, to untutored ears, as if strict implication is indeed far removed from ordinary implication. The fact is, however, that he used the name "strict implication" in order to differentiate between the kind of implication which functions in his systems, and which he thought to be *identical* with ordinary implication, and the relation for which Whitehead and Russell, in their epoch-making *Principia Mathematica* (1910–13), had co-opted the expression "implication" (nowadays called "material implication"). In any case, the kind of implication which features in his systems has it in common with ordinary implication that in both cases it is a necessary condition of P implying Q that it should *not be possible* for P to be true and Q to be false. By way of contrast, the "implication" of *Principia Mathematica* was a merely truth-functional relation which held between P and Q whenever *as a matter of fact* it is not the case that P is true when Q is false. It was *this* account of implication, Lewis felt, which was artificial, specialized, and purely formal. The use of the word "strict", in his characterization of the kind of implication found in his systems, was introduced to effect a contrast with Whitehead and Russell's use of the term "implication". In short, his use of the term "strict" in "strict implication" was occasioned by a mere quirk of history. It needs to be remembered, then, that when Lewis defined strict implication as the relation which holds between P and Q when it is not *possible* that P be true and Q be false, he took himself to be defining ordinary implication.[60] Interestingly enough, ancient and medieval logicians had taken themselves to be doing exactly the same thing when they, too, had defined implication in the same way. Thus, for instance, we find that in the fourth century B.C., Diodorus Cronus offered the same modal analysis as Lewis and contrasted it with what Diodorus regarded as the eviscerated truth-functional account offered by his gifted pupil, Philo of Megara. Plainly, the dispute about implication had a precursor twenty-four centuries before Lewis took issue with Whitehead and Russell. And it also had a precursor in the medieval period when the Schoolmen contrasted the very same modal account as Lewis gave with the truth-functional one of Philo.

That Lewis was right about the *necessary* conditions for P being said, in the ordinary sense, to "imply" Q, is seldom disputed. But was he right in claiming further that the impossibility of P being true while Q is false, is a *sufficient* condition of P implying Q? This is where we encounter the objection based on the "paradoxes" T10 and T11. For T10 and T11 are generated only if we take the impossibility of P being true and Q false to be a sufficient condition.

Lewis' main reply was that the so-called paradoxes are not really paradoxical at all and that, to see that they are not, we need only reflect on the fact that our ordinary intuitions about what implication is, and about what implies what, commit us to them. He therefore gave two *independent* proofs of theses which are special cases of T10 and T11, respectively — proofs which do not depend

60. See C.I. Lewis, *A Survey of Symbolic Logic,* Berkeley, University of California Press, 1918, esp. p. 324.

upon any alleged artificialities in his axiomatic systems S2 to S5, but depend only upon what we would ordinary agree to be valid rules of inference in any context or sphere of discourse. We give the first proof (for a special case of T10) only. It requires merely that we subscribe to each of the inference rules, Simplification, Addition, and Disjunctive Syllogism.

Consider, first, some necessarily false proposition which asserts

(a) $P \cdot \sim P$

From (a), by Simplification, we may validly infer

(b) P

From (b), by Addition, we may validly infer

(c) P or Q

From (a), by Simplification, we may validly infer

(d) $\sim P$

But from (c) and (d), by Disjunctive Syllogism, we may validly infer

(e) Q

where Q, of course, may be any proposition at all. It is clear that by a series of steps — each warranted in the manner of natural deduction systems by a rule of valid inference — we can deduce any proposition whatever from a necessarily false proposition of the form $P \cdot \sim P$. And since, as we saw earlier, an inference (or series of inferences) from one proposition to another is valid if and only if the former *implies* the latter, we have here a proof that a necessarily false proposition of the form $P \cdot \sim P$ implies any proposition whatever. That is, we have here a proof of a special case of the so-called "paradox" T10.[61]

It is sometimes suggested that "all one has to do" in order to avoid the conclusion of this independent argument is to give up one of the rules of inference cited. But the situation is nowhere near as simple as that. The conclusion can be avoided only at the cost of giving up *at least one more* of the ordinarily accepted rules of inference. For, without using *any* of the above-mentioned rules, we can give *another* proof for a special case of T10 (viz., for $[P \cdot \sim P] \rightarrow \sim Q$), and a proof for a special case of T11. The proof of the former requires only that we agree to a proposition which is ordinarily agreed to state a fact of implication, viz.,

T16: $(P \cdot Q) \rightarrow P$[62]

61. The proof is a proof of $(P \cdot \sim P) \rightarrow Q$. We call this a "special case" of T10 because it does not have the full generality of T10. Thus T10 does not require that the necessarily false (impossible) proposition from which any proposition follows be in the form of an explicit contradiction. It claims that *any* necessarily false proposition — whether or not of this form — implies any proposition. (Similar remarks apply, mutatis mutandis, to our remarks about the special case that is provable for T11.)

62. $(P \cdot Q) \rightarrow P$, of course, is the 'fact' about implication which corresponds to the Rule of Simplification.

and agree, further, to a rule of inference which, ever since Aristotle, has been recognized as valid, viz.,

Antilogism:

> From the claim that two propositions imply a third we may validly infer that either of them together with the negation of the third implies the negation of the other.

From T16, in one step, by Antilogism, we may make a valid immediate inference to

T17: $(P \cdot \sim P) \rightarrow \sim Q$

Yet T17 is simply a special case of T10 since it asserts that any necessarily false proposition of the form $P \cdot \sim P$ implies any proposition whatever.[63]

The other proof requires only that we accept the proof just given as valid and agree, further, to the already familiar rule of inference known as

Transposition:

> From the claim that one proposition implies another, one may validly infer that the negation of the latter implies the negation of the former.

From T17, in one step, by Transposition, we may validly infer

T18: $Q \rightarrow \sim (P \cdot \sim P)$

Yet T18 is simply a special case of T11 since it asserts that any necessarily true proposition of the form $\sim (P \cdot \sim P)$ [necessarily true because it asserts the negation of the necessarily false $(P \cdot \sim P)$] is implied by any proposition whatever. Clearly, in order to avoid the "paradoxical" T17 we should have to give up either T16 or the rule of Antilogism. And in order to avoid the "paradoxical" T18 we should have to give up at least one of these or the rule of Transposition. What these two proofs, together with those of Lewis, show is that cases of the alleged paradoxes can be avoided only at the cost of *more than one* of our ordinary intuitions about what implies what and what may validly be inferred from what.

63. As a matter of fact, this proof can be extended to establish that *every* necessarily false proposition — not only those of the form "$(P \cdot \sim P)$" — implies every proposition whatever. One need only invoke the truth that every necessarily false proposition implies both itself and its negation, i.e., the principle $[\Box \sim P \rightarrow (P \rightarrow [P \cdot \sim P])]$, and the proof — using the method of mediate conditional proof (abbreviated "C.P.") of natural deduction — is straightforward:

(1)	$\Box \sim P \rightarrow (P \rightarrow [P \cdot \sim P])$	
(2)	$(P \cdot \sim P) \rightarrow \sim Q$	[T17]
(3)	$\Box \sim P$	[Assumption]
(4)	$P \rightarrow (P \cdot \sim P)$	[(1), (3) × Modus Ponens]
(5)	$P \rightarrow \sim Q$	[(4), (2) × Hypothetical Syllogism]
(6)	$\Box \sim P \rightarrow (P \rightarrow \sim Q)$	[(3)-(5) × C.P.] Q.E.D.

EXERCISE

It is sometimes supposed that the 'paradoxes' of strict implication, i.e., the proposition that a necessarily true proposition is implied by any proposition and the proposition that a necessarily false proposition implies any proposition, may be expressed in this way:

"$Q \rightarrow \Box P$";

"$\Box \sim P \rightarrow Q$".

Try to explain why these are incorrect *paraphrases of the prose claims and do* not *express those propositions.*

* * * * *

Relevance logics

Now some logicians say they are prepared to pay this cost in order to avoid the paradoxes. As proof of their willingness, some have actually constructed deductive systems within which many of our ordinary intuitions about implication are preserved but the paradoxical theorems are not. For instance, the System E of Alan Ross Anderson and Nuel D. Belnap[64] manages to avoid the paradoxes at the cost of rejecting the rule of Disjunctive Syllogism. And a good many other logicians have worked at constructing so-called *Relevance Logics* which it is hoped will achieve the same end by other similar means. Their dissatisfaction with any account (such as Lewis') which holds it to be a sufficient as well as necessary condition of P implying Q that it should be impossible for P to be true and Q false, plainly runs very deep. It cannot be dismissed as stemming from any superficial misunderstanding of Lewis' term "strict implication". Neither does it stem merely from reaction to the unexpectedness of the consequences (T10 and T11) which his account generates.[65] It stems rather from a deep conviction that *more* is required for the relation of implication to hold between P and Q, where the "more" is seen as being some "inner connection", some "identity of content", or "connection of meaning" between P and Q. This is what is meant when, in discussions of Relevance Logics, it is said that *relevance* is also a necessary condition of P implying Q. The complaint leveled against the Lewis-type definition of implication is that it commits us to the view that implication can hold between two propositions in a purely "external" way. The paradoxes, it would be said, are merely *symptoms* of the defect to which they are reacting: they are not the principal defect itself.

A detailed examination of the pros and cons of Relevance Logics cannot be undertaken here. We will venture just a few brief remarks to help set the issue into perspective.

What Relevance Logicians are getting at can be made clear if we consider a particular substitution-instance of one of the paradoxes; e.g., T11: $\Box P \rightarrow (Q \rightarrow P)$. Let P be the proposition

(4.27) $9 = 3^2$

and let Q be the proposition

64. Anderson and Belnap, "The Pure Calculus of Entailment", *Journal of Symbolic Logic*, vol. 27 (March 1962) pp. 19–52.

65. If it did, Lewis' claim that they are "paradoxical only in the sense of expressing logical truths which are easily overlooked" should suffice as an answer. See C.I. Lewis and C.M. Langford, *Symbolic Logic*, The Century Co., 1932; second edition, New York, Dover, 1959, p. 248.

(4.28) The mists are hanging low today.

Let us agree that *(4.27)* is necessarily true; i.e., let us assume the truth of

(4.29) $\Box(9 = 3^2)$

Then, since the rule of inference corresponding to T11 says that if a proposition is necessarily true we may validly infer that it is implied by any proposition whatever, we may validly infer from *(4.28)* the proposition

(4.30) (The mists are hanging low today)$\rightarrow(9 = 3^2)$

But *(4.30)*, the Relevance Logician points out, is counterintuitive. And it is counterintuitive, he further tells us, just because the truth of *(4.28)* is *irrelevant* to the truth of *(4.27)*. *This,* he concludes, is what is paradoxical about the Lewis-type definitions: they lead us to hold that propositions imply one another when the relevance-condition is not satisfied.

Their complaint is obviously connected closely with that which we discussed, in chapter 1, about the paradoxicality of the claim that any two necessarily true propositions are logically equivalent to one another. That claim, it should be obvious, is precisely what is asserted in systems S2 to S5 by the thesis

T12: $(\Box P \cdot \Box Q)\rightarrow(P\leftrightarrow Q)$

Its apparent paradoxicality, it will be remembered, stemmed from our disinclination to say of the necessarily true propositions

(1.5) Either the U.S. entered World War I in 1917 or it did not

and

(1.23) Either Canada is south of Mexico or it is not

that they are *really* equivalent even though our definition of "equivalence" forced this conclusion upon us. They have nothing to do with one another, we were inclined to say; so how *could* they be equivalent? It seems clear in retrospect that the qualms thus expressed about equivalence were rooted in the same sort of qualms which Relevance Logicians have expressed about implication. If two propositions have nothing to do with one another, how can one imply the other (as is claimed by T10 and T11) let alone be equivalent to the other (as claimed by T12)?

Our way of handling the earlier problem about equivalence suggests a way of dealing with, or at least of throwing some light on, the related problems about implication. We then suggested that the air of paradox involved in the claim that *(1.5)* is equivalent to *(1.23)* could be removed by recognizing that propositional identity is merely a special case of propositional equivalence so that, although all cases where an identity-relation obtains will be cases where an equivalence-relation obtains, we should not expect the converse to hold. We have a similar suggestion to offer about implication. Let us allow that whenever certain sorts of "inner connection" or "identity of content" obtain between two propositions, as in the cases of proposition-pairs such as

(4.31) Pat is someone's sister

and

(4.32) Pat is female

or again

(4.16) If it rains the snow will melt, and if the snow melts the World Cup slalom will be cancelled

and

(4.17) If it rains, the World Cup slalom will be cancelled

then the relation of implication will obtain between these propositions. But let us not expect that the converse will also hold in every case. In other words, we suggest that although the finding of the right sort of identity of content is sufficient ground for concluding that an implication relation holds, it is a mistake to suppose it also to be necessary.

It is easy to see how the demand for relevance or inner connection arises. The relevance-condition is *automatically* satisfied in so many of the 'ordinary' cases of implication that come before us: it is satisfied in all cases of implication between contingently applicable concepts; it is satisfied in all cases of implication relations between contingent propositions;[66] and it is even satisfied in many cases of implication relations between noncontingent concepts as well as propositions. Little wonder, then, that the expectation is generated that the relevance condition should be satisfied in *every* case of implication, and a fortiori in *every* case of equivalence.

It is tempting to dismiss the Relevance Logicians' demands for inner connections, and their consequent criticisms of systems like S5 (which accept the impossibility of P being true and Q false as a sufficient condition for P implying Q), by saying that these demands and criticisms are to be attributed to what Wittgenstein once called "a main cause of philosophical disease", viz., "a one-sided diet: one nourishes one's thinking with only one kind of example" (*Philosophical Investigations,* § 593). But that would be too cavalier. For although the proponents of Relevance Logic seem, up to this point, to have had little or no success in defining the concept of *relevant implication*, there can be little doubt but that such a concept is worth defining. We can characterize such a concept broadly by saying that it has application to a proper subset of the cases in which the relation of strict or logical implication holds; and we can say that it stands to the concept of logical implication (definable in terms of possible worlds) in much the same sort of way as the concept of propositional identity stands to the concept of propositional equivalence. The difficulty is to characterize the finer-grained concept of relevant implication more precisely than that.

One thing is clear, however. Nothing is gained by saying that strict or logical implication isn't *really* a case of implication at all. And nothing is gained — but on the contrary much is lost — by insisting that certain standard rules of inference which intuitively strike us as valid are really not so. In rejecting, as invalid, rules such as those of Addition and Disjunctive Syllogism we would be committed also to rejecting each of the analytical methods whereby those rules are customarily justified.[67] And if we give up these, we seem left with no recourse to reason as a way of providing backing for any of our logical intuitions. The cost, in brief, seems prohibitive — prohibitive of reason itself.

The move to predicate logic

Whether or not the following of one proposition from another is *always* dependent (as Relevance Logicians believe) upon the existence of an internal connection between them, there can be no doubt

66. Recall our suggestion (in chapter 1, p. 54) that much of the air of paradox generated by possible-worlds analyses of implication is due to a preoccupation with contingent propositions.

67. Among the analytical methods that would have to be abandoned are those of truth-table analysis, given in chapter 5, and the reductio technique, given in chapters 5 and 6. Needless to say, our use of worlds-diagrams as a decision-procedure for truth-functional and modal propositional logic — demonstrated in chapters 5 and 6 — would also go by the board.

that our ability to *show* that one proposition follows from another is *often* dependent upon our ability to *show* that there is a certain sort of internal connection between them. Analysis of the kind which is achieved within the logic of propositions — the logic of unanalyzed propositions, that is — may fail to show that one proposition follows from (or, conversely, implies) another just because it neglects all matters to do with the internal structure of simple propositions and hence neglects all matters to do with the internal connections between simple propositions.

To be sure, propositional logics do not neglect the internal structure of *compound* propositions such as

(4.1) Either it is necessarily true that sisters are female or it is not

and

(4.2) It is necessarily true that sisters are female.

Truth-functional propositional logic can tell us that *(4.1)* is to be analyzed as having the structure of a compound proposition which is the disjunction of two contradictories; and it will exhibit this structure by saying that *(4.1)* has the form $P \vee \sim P$. Likewise, modal propositional logic can tell us that *(4.2)* is to be analyzed as having the structure of a compound proposition which ascribes necessary truth to a simpler proposition; and it will exhibit this structure by saying that *(4.2)* has the form $\Box P$. Moreover, by virtue of thus analyzing these compound propositions, these two kinds of propositional logic can reveal a great deal about the internal (logical) connections between these compound propositions and other propositions. But neither logic tells us anything about the internal structure of the *simple* propositions which they involve. And as a consequence neither can tell us anything about any internal, logical, connections which these simple propositions may bear to one another.

In order to show that certain propositions imply others, or that certain corresponding inferences are valid, we often (though not, of course always) must take recourse to such details of the structure of simple propositions as is revealed by analysis at a deeper level: that provided by the logic of predicates — of unanalyzed concepts, that is.

Consider, for example, the argument from the propositions

(4.33) All politically enlightened persons are sympathetic to socialism

and

(4.34) All women's liberationists are politically enlightened

to the proposition

(4.35) All women's liberationists are sympathetic to socialism.

No matter what one thinks of the truth of either of the premises or of the conclusion of this argument, there can be no doubt of its *validity:* the conclusion follows from the premises; the premises imply the conclusion. But how can this be *shown?* Here the logic of propositions cannot help us. By employing the meager analytical and notational resources of propositional logics we can show that the argument has a certain structure or form, one which we might record by writing

$$\begin{array}{ll} & P \\ & Q \\ \hline \therefore & R \end{array}$$

But there is nothing about the analyzed structure of the argument as thus exhibited which entitles us to conclude that the argument is valid. The argument

(4.34)		All women's liberationists are politically enlightened
(4.36)		All persons sympathetic to socialism are politically enlightened
(4.37)	∴	All persons sympathetic to socialism are women's liberationists

has precisely the same form, as revealed at that level of analysis, and yet it is patently *invalid.* After all, there are probably many persons who believe both *(4.34)* and *(4.36)* to be true and yet would strenuously deny *(4.37);* and even if their beliefs are mistaken, it is clear that they cannot fairly be charged with inconsistency as can anyone who asserts the premises but denies the conclusion of a *valid* argument.

Nevertheless, the validity of the first argument can be shown. It can be shown once we employ the richer analytical and notational resources of predicate logic.

Traditional syllogistic

Aristotle was the first known logician to put any of these requisite resources at our disposal. He put the science of formal logic on its feet by formulating the rules whereby the validity of arguments of this sort — *syllogistic* arguments, as they are called — may be determined. Within the traditional syllogistic logic which he established, the first argument may be analyzed as having the form:

All M are P
All S are M
∴ All S are P

and the second as having the form

All P are M
All S are M
∴ All S are P

where the form of the argument is determined by (a) the internal structure of each proposition and (b) the connection between these internal structures. As to (a), traditional logic analyzed propositions (of the kind that occur in syllogisms) as having one or other of four possible forms:

A:	All . . . are . . .	(Universal affirmative)
E:	No . . . are . . .	(Universal negative)
I:	Some . . . are . . .	(Particular affirmative)
O:	Some . . . are not . . .	(Particular negative)

where the blanks are filled by so-called "terms". (A term may be regarded, from the point of view of modern logic, as an expression which stands for a property which can be predicated of some item or other, e.g., being a women's liberationist, being politically enlightened, being sympathetic to socialism, etc.) As to (b), traditional logic recognized that the terms which occur within the premises and conclusion can occur within the *argument* in one or other of four possible ways known as "*Figures* of Syllogism". Thus, where "S" stands for the term which occurs as *subject* of the conclusion as well as in one of the premises, "P" stands for the term which occurs as *predicate* of the conclusion as well as in one of the premises, and "M" stands for the so-called *middle term*, i.e., the term which occurs twice in the premises, the four figures of syllogism are:

$$\begin{array}{ccccccccccc}
 & \mathrm{I} & & & \mathrm{II} & & & \mathrm{III} & & & \mathrm{IV} \\
 & \mathrm{M}\quad\mathrm{P} & & & \mathrm{P}\quad\mathrm{M} & & & \mathrm{M}\quad\mathrm{P} & & & \mathrm{P}\quad\mathrm{M} \\
 & \underline{\mathrm{S}\quad\mathrm{M}} & & & \underline{\mathrm{S}\quad\mathrm{M}} & & & \underline{\mathrm{M}\quad\mathrm{S}} & & & \underline{\mathrm{M}\quad\mathrm{S}} \\
\therefore & \mathrm{S}\quad\mathrm{P} & & \therefore & \mathrm{S}\quad\mathrm{P} & & \therefore & \mathrm{S}\quad\mathrm{P} & & \therefore & \mathrm{S}\quad\mathrm{P}
\end{array}$$

It can easily be seen that, provided one abides by the convention of always writing the premise which contains the predicate of the conclusion (the *major term*, as it is called) first, and the premise which contains the subject of the conclusion (the *minor term*, it is called) second, these are the only ways in which the major, minor, and middle terms can occur. It can also be seen that since, on this analysis, each of the three propositions involved in each figure may itself have any one of four internal structures or forms (those cited in (a)), there are altogether 4×4^3 (= 256) distinct ways in which the internal structures of the propositions in a syllogism may be connected. Thus there are 256 possible *forms of syllogism*. Needless to say, of the 256 only a relative handful exhibit modes of connection all of whose instances are valid arguments. One of Aristotle's great achievements was to list all the valid forms and provide a set of rules by means of which to distinguish them from the others.

The details of traditional syllogistic analysis need not concern us here. The main point to note is that the validity of many arguments can be determined simply by analyzing them to the level made possible within that tradition and checking to see whether the form which, on analysis, that argument is found to have, is one of the certifiably valid ones. *We do not have to analyze the terms themselves or even understand what concepts they express* in order to show that certain arguments are valid. For instance, the form of the argument from the conjunction of *(4.33)* and *(4.34)* to *(4.35)* turns out to be one of the valid ones. By way of contrast, the form of the argument from the conjunction of *(4.34)* and *(4.36)* to *(4.37)* turns out *not* to be one of the valid ones. Note that we do not say that any argument whose form is not certifiably valid is an argument which can be certified as invalid. Plainly, that would be a mistake. As we have already seen, in passing from the logic of unanalyzed propositions to the logic of unanalyzed concepts, an argument whose form at one level of analysis is not certifiably valid may turn out, at a deeper level of analysis, to be valid nonetheless. Having a certifiably valid form is a sufficient condition of the validity of an argument but it is not a necessary condition. (We will make more of this point in chapter 5.)

Modern predicate logic

The analytical and notational resources of modern predicate logic are much richer than those of traditional syllogistic. Accordingly, many more arguments yield to its treatment. Like traditional syllogistic, it recognizes that so-called *quantifier-words*, like "all" and "some", express concepts which feature in the internal structure of a proposition in such a way as to determine that proposition's logical connections with other propositions independently of what other concepts feature in that proposition, and independently, too, of the analysis of those concepts. But unlike traditional syllogistic it utilizes a symbolism which blends with that of propositional logic to provide a much more versatile means of exhibiting the internal structure of propositions. Consider, for instance, the way in which modern predicate logic enables us to validate the argument from *(4.33)* and *(4.34)* to *(4.35)*.

First let us analyze the propositions themselves. The proposition

(4.33) All politically enlightened persons are sympathetic to socialism

is analyzed as asserting that if any items have the property of being a politically enlightened person then those items have the property of being sympathetic to socialism. Using the individual variable "x" to stand for any item whatever,[68] and the predicate letters "**P**" and "**S**" to stand for the properties of being politically enlightened and being sympathetic to socialism, we can then render this analysis in symbols as

$$(x)\ (\mathbf{P}x \supset \mathbf{S}x)$$

to be read as "For any x, if x has the property **P** then x has the property **S**". Similarly, the proposition

(4.34) All women's liberationists are politically enlightened

is analyzed as having the form

$$(x)\ (\mathbf{W}x \supset \mathbf{P}x)$$

where "**W**x" bears the obvious interpretation "x has the property of being a women's liberationist". And the conclusion

(4.35) All women's liberationists are sympathetic to socialism

is analyzed as having the form

$$(x)\ (\mathbf{W}x \supset \mathbf{S}x).$$

The argument can then be set out thus:

$$\begin{array}{rl} & (x)\ (\mathbf{P}x \supset \mathbf{S}x) \\ & \underline{(x)\ (\mathbf{W}x \supset \mathbf{P}x)} \\ \therefore & (x)\ (\mathbf{W}x \supset \mathbf{S}x) \end{array}$$

68. The lowercase letters at the end of the alphabet are standardly used as *individual variables*, i.e., so as to refer indiscriminately to any individuals whatever. Since individual variables are used not to refer to any particular items but indiscriminately to any items whatever, it follows that two or more distinct variables (e.g. "x" and "y") may have one and the same item as their referents.

Now we could, at this point, simply appeal to the educated logical intuitions of anyone who understands the symbolism to validate the argument. And for the purpose of this exercise it would obviously not matter whether one understood what the predicate letters "**P**", "**S**", or "**W**" stood for (what properties they denoted or what concepts they expressed). But intuition is not always reliable. And in any case, the rich resources of predicate logic are at hand.

In order to *show* that the argument is valid we appeal to the already familiar rule of Hypothetical Syllogism along with two rules which belong to predicate logic, viz., the rules of Universal Instantiation (U.I.) and Universal Generalization (U.G.). U.I. tells us that whatever is true of every item is also true of any given item. And U.G. in effect tells us that we can infer a truth about every item from a truth about an arbitrarily selected item.[69] We can then prove that the conclusion follows from the premises by constructing a series of steps from premises to conclusion, each step being justified by appeal to a valid rule of inference. The proof goes as follows:

(1)	$(x)\ (\mathbf{P}x \supset \mathbf{S}x)$	[Premise]
(2)	$(x)\ (\mathbf{W}x \supset \mathbf{P}x)$	[Premise]
(3)	$\mathbf{P}x \supset \mathbf{S}x$	[(1) × U.I.]
(4)	$\mathbf{W}x \supset \mathbf{P}x$	[(2) × U.I.]
(5)	$\mathbf{W}x \supset \mathbf{S}x$	[(4), (3) × Hypothetical Syllogism]
(6)	$(x)\ (\mathbf{W}x \supset \mathbf{S}x)$	[(5) × U.G.]

The proof offered is constructed in the style of natural deduction rather than axiomatics. But needless to say, a proof of the validity of the argument could equally well, though with a good deal more difficulty, be given in an axiomatization of predicate logic, i.e., in the so-called Predicate Calculus.

So far, the only quantifier we have used is the universal quantifier "(x)". With its help we can analyze and render into symbolic form propositions which make assertions about *all* items having a certain property. But not all propositions make universal claims. Sometimes we merely want to make the lesser claim that there is *at least one* item which has a certain property. In order to give straightforward expression to such claims, modern predicate logic uses the *existential quantifier* "$(\exists x)$" — to be read as "There is at least one item such that. . . . " Thus if we wanted to analyze the proposition

(4.32) Some persons sympathetic to socialism are women's liberationists

we could write

$$(\exists x)\ (\mathbf{S}x \cdot \mathbf{W}x)$$

and read it as "There is at least one x such that x is an **S** and x is a **W**". Strictly speaking, any formula containing an existential quantifier can be rewritten in the form of one containing a universal quantifier, and vice versa, as the following equivalences make clear:

69. More perspicuously, U.G. could be stated this way: If a property holds of a member of a set irrespective of which member it is, then that property holds of every member of that set.

$$\text{“}(x)\,(\mathbf{S}x \supset \mathbf{W}x)\text{”} =_{df} \text{“}{\sim}(\exists x)\,(\mathbf{S}x \cdot {\sim}\mathbf{W}x)\text{”}$$

$$\text{“}(\exists x)\,(\mathbf{S}x \cdot \mathbf{W}x)\text{”} =_{df} \text{“}{\sim}(x)\,(\mathbf{S}x \supset {\sim}\mathbf{W}x)\text{”}$$

Nevertheless, the symbolism is a lot easier to read and to work with if we allow this small redundancy.

Modal notions in predicate logic

Our purposes in giving the foregoing sketches of the symbolism of traditional and modern predicate logic have been twofold. First, we have wanted to illustrate the fact that in the case of many valid inferences, the analytical and notational resources of *propositional* logic do not suffice to show us why these inferences are valid. For such cases, we need a deeper analysis such as that provided by *predicate* logic. Secondly, we have wanted to prepare the ground for an intelligible discussion of the role which modal concepts play at this deeper level of analysis. We are now ready for that discussion.

Although the object-languages of traditional syllogistic and modern predicate calculus contain no symbols for the modal properties of necessity, possibility, etc., or the modal relations of implication, equivalence, etc., it is clear that insofar as these systems are taken to establish the logical truth of certain theses or the validity of certain argument-forms, modal concepts are *implicitly* invoked.

They are *explicitly* invoked in modal predicate logic. Once more, it was Aristotle who did the pioneering work. His treatment of modal syllogistic, it has been conjectured, was his last major contribution to logic. That treatment, however, although it motivated much medieval interest in modal concepts, was far from satisfactory. It was not until the 1940s that Ruth C. Barcan (later Ruth Barcan Marcus) investigated ways of blending modal logic with modern theory about the quantifiers "(x)" and "(∃x)" and so founded modal predicate logic as it is usually understood.[70]

Many of the most interesting, and also many of the most controversial, questions about modal predicate logic concern the formula by means of which Barcan tried to effect the 'mixing' of the two kinds of logic. The formula, which has come to be known as "the Barcan Formula" (BF) may be symbolized in two forms:

$$\text{BF1:}\quad \Diamond(\exists x)\mathbf{F}x \rightarrow (\exists x)\Diamond\mathbf{F}x$$

(the form in which she originally propounded it); or as

$$\text{BF2:}\quad (x)\,\Box\mathbf{F}x \rightarrow \Box(x)\mathbf{F}x$$

(which can easily be shown to be equivalent to BF1, and which is the form which has most often attracted attention). BF1 may be read as asserting: From the proposition that it is possible that there exists an item which has the property **F** it follows that there exists an item which possibly has **F**. And BF2 may be read as asserting: From the proposition that every item necessarily has the property **F** it follows that it is necessary that everything has the property **F**.

There is nothing surprising nor controversial about the presence in BF1 and BF2, respectively, of the wffs $\Diamond(\exists x)\mathbf{F}x$ and $\Box(x)\mathbf{F}x$. The wff $\Diamond(\exists x)\mathbf{F}x$ is easily recognizable as a substitution-instance of the wff $\Diamond P$ of modal propositional logic; and $\Box(x)\mathbf{F}x$ is easily recognizable as a substitution-instance of the wff $\Box P$ of modal propositional logic. The surprises, and the puzzles, are to be found in the rest of each of these formulae: (a) in the mere presence of the wffs $(\exists x)\Diamond\mathbf{F}x$ and $(x)\Box\mathbf{F}x$ in BF1 and BF2, respectively; and (b) in the asserted implications whereby $(\exists x)\Diamond\mathbf{F}x$ is claimed to follow from $\Diamond(\exists x)\mathbf{F}x$ and $(x)\Box\mathbf{F}x$ is claimed to imply $\Box(x)\mathbf{F}x$.

70. Independently and concurrently, Carnap developed foundations for modal predicate logic.

Modalities de dicto and de re

The mere presence of the wffs $(\exists x)\Diamond \mathbf{F}x$ and $(x)\Box \mathbf{F}x$ seems to some philosophers to be philosophically suspect, for it reminds them of a distinction which medieval logicians made much of: that between modalities *de dicto* and modalities *de re*. By a *de dicto* modality, as Thomas Aquinas explained it, is meant the attribution of a modal property to a *proposition* as in the proposition

(4.39) It is possible that Socrates is running

whereas by a *de re* modality is meant the attribution of a modal property to an *individual* as in the proposition

(4.40) Socrates is possibly running.

The distinction itself, it would be admitted, is not particularly troublesome; indeed it reflects accurately enough the two main uses of modal expressions in natural languages such as Latin and English. What is troublesome, they would say, is what some philosophers have *said* about the distinction. Some philosophers have said that *de re* modalities are irreducibly different from *de dicto* ones and that, accordingly, it makes sense to revive the Aristotelian doctrine of *essentialism,* i.e., the doctrine that some properties inhere essentially or necessarily in the individuals which have those properties. As against this, many philosophers, such as Quine, regard essentialism as an anachronism which deserves no place in a scientific view of the world. Accordingly, philosophers of Quine's conviction find expressions like $(x)\Box \mathbf{F}x$ and $(\exists x)\Diamond \mathbf{F}x$, in which the modal symbols appear in *de re* position, thoroughly misleading. If these wffs are merely notational variants on the corresponding wffs *de dicto,* viz., $\Box(x)\mathbf{F}x$ and $\Diamond(\exists x)\mathbf{F}x$, then — they would say — quantified modal logic is an unnecessary complication. But if they are taken to be irreducibly different from their *de dicto* counterparts, then — they would say — quantified modal logic is metaphysically objectionable.

This first objection to BF1 and BF2 can, of course, be met by arguing that there is nothing at all wrong with essentialism and that the contrary view, expressed by John Stuart Mill in the words "Individuals have no essences", is itself insupportable. There has in fact been a revival in recent years of interest in, and support for, the doctrine of essentialism. Unfortunately we cannot pursue the issue here.[71]

A second objection to BF1 and BF2 is that each of the asserted implications seems to be exposed to obvious counterexamples. As an instance of

$$\text{BF1:}\quad \Diamond(\exists x)\mathbf{F}x \rightarrow (\exists x)\Diamond \mathbf{F}x$$

consider the case where $\mathbf{F}x$ expresses the concept of being someone who landed in Kansas in 1916 from a space-yacht called "Dora". Then BF1 commits us to saying that the proposition

(4.41) It is possible that there exists someone who landed in Kansas in 1916 from a space-yacht called "Dora"

implies the proposition

(4.42) There exists someone who possibly landed in Kansas in 1916 from a space-yacht called "Dora".

71. For a readable defence of essentialism and discussion of some of the issues surrounding it, see Baruch A. Brody, "Why Settle for Anything Less than Good Old-Fashioned Aristotelian Essentialism?" *Nous,* vol. 7 (1973), pp. 351–365.

Suppose, however, that what makes *(4.41)* true is the fact that in the possible world of Heinlein's novel *Time Enough For Love* the chief character, Lazarus Long, has the property of landing from a space-yacht, etc. Does it follow from this that there really *is* someone (someone in the actual world) who possibly has that property? Surely not. Although Lazarus may exist in the possible world of *Time Enough For Love,* he may well not exist in the actual world — and, for that matter, neither may anyone else of whom it is true to say that he might have landed in a space-yacht in 1916. *(4.42)* cannot follow from *(4.41)* since it may be false when *(4.41)* is true. Again, as an instance of

BF2: $(x) \Box \mathbf{F}x \rightarrow \Box (x)\mathbf{F}x$

consider the case where $\mathbf{F}x$ expresses the concept of being something that exists. Then BF2 commits us to saying that the proposition

(4.43) Everything necessarily exists

implies the proposition

(4.44) It is necessarily true that everything exists.

Suppose, however, that we hold *(4.43)* to be true because, like some essentialists, we hold that existence is an essential property of everything that *actually* exists.[72] Does it follow from this that, as *(4.44)* asserts, in the case of every possible world everything that exists in the actual world exists there also? Hardly. Although Nixon exists in the actual world and hence, according to some essentialists, essentially exists therein, he surely does not exist in all the possible but nonactual worlds which, in our more fanciful moments, we conceive of.

All this is very puzzling. Not only have many astute thinkers accepted the Barcan formulae as obvious truths of logic; it also turns out that these formulae are derivable as theorems in certain axiomatizations of modal predicate logic, viz., in certain axiom systems which combine the truth-functional predicate calculus with S5. On the face of it, then, if we were to accept the purported counterexamples given above then we should have to reject either the truth-functional predicate calculus or the modal system S5. Neither seems a palatable alternative. But, then, too, the counterexamples to the Barcan formulae also seem very persuasive.

There is a way out of this logical bind. It turns out that the Barcan formulae are not derivable in *all* axiomatizations of predicate logic but only in some. They are derivable only in axiomatizations which yield, as theorems, formulae some of which contain what are called *free variables.* They are not derivable in axiomatizations which yield, as theorems, formulae all of which are said to be universally closed.[73] This means that the choice before us is not quite as painful as it might have seemed. We can continue to accept the counterexamples as genuine, and continue to accept S5, simply by deciding to accept as theorems only those formulae of predicate logic which are universally closed. Quantified S5, thus presented, does not contain either of the Barcan formulae.[74]

72. Brody, *op. cit.,* holds this. **F** is an essential property of an object, O, on his view, just when O has that property and would go out of existence if it lost it. Since nothing can continue to exist if it loses the property of existence, existence — on his view — is an essential property. On his view, that is, *(4.43)* is true.

73. Roughly, a formula contains a free variable if it contains a variable which is not subject to quantification. Thus, in $(x)(\mathbf{F}x \supset \mathbf{G}y)$, y is free since it is not "bound" by, or subject to, the quantifier (x). The formula $(x)(\mathbf{F}x \supset \mathbf{G}y)$ is not universally closed. We can make it into a universally closed formula, however, by subjecting the free variable y to universal quantification as in $(y)(x)(\mathbf{F}x \supset \mathbf{G}y)$.

74. This important result was first proved by Saul Kripke, "Semantical Considerations on Modal Logic", *Acta Philosophica Fennica,* vol. 16 (1963), especially pp. 57–90. As Kripke points out, the acceptance, as theorems, of

We are still left with the puzzle that, quite independently of the alleged derivability of the Barcan formulae within quantified S5, many philosophers have found these formulae intuitively acceptable. Why should this be? We might be tempted, at this point, to invoke the hypothesis that these philosophers simply have not subjected their beliefs in the Barcan formulae to that kind of strenuous search for counterexamples which, in chapter 2, we described as the Method of Possible-Worlds Parables.[75] If so, we are inclined to say, they would surely have turned up the counterexamples cited. and accordingly have abandoned these beliefs. But this hypothesis would not do justice to the situation. The fact is that the sponsors of the Barcan formulae accept these formulae not out of ignorance of the existence of purported counterexamples, but because they have a view of what possible worlds are which does not allow us, without inconsistency, even to construct these supposed counterexamples. Let us explain.

Heterogeneous and homogeneous possible worlds

It is clear, on reflection, that in offering these counterexamples we were presupposing that an object which exists in one possible world might not exist in another. Thus our counterexample to BF1 depended upon the assumption that Lazarus Long exists in some nonactual possible worlds even though he does not exist in the actual one. And our counterexample to BF2 depended upon the assumption that even though Nixon does exist in the actual world he does not exist in some non-actual possible world. In other words, we have been supposing that possible worlds are *heterogeneous* in respect of which objects they contain: that some objects which do not exist in the actual world do exist in other possible worlds, and that some objects which do exist in the actual world do not exist in other possible worlds.

But suppose we were to take the view that possible worlds are *homogeneous* in respect of which objects they contain: that all and only those objects which exist in the actual world can intelligibly be supposed to exist in other possible worlds. Then other possible worlds will differ from ours only in respect of the differing properties which these objects have and in respect of the differing relations in which these objects stand to one another. But since Lazarus does not exist in the actual world, there will not be any possible worlds in which he does exist; and since Nixon exists in the actual world, there will not be any possible worlds in which he fails to exist. On this *homogeneous-worlds* view, the counterexamples simply cannot be envisaged and the Barcan formulae express obvious truths. [76]

At this point it is tempting to ask: Which of these views about possible worlds is the correct one? Tempting, perhaps; but not a question to be pursued here. Our own view of the matter should be evident from the fact that we chose to introduce possible worlds, in chapter 1, p. 1, by reference to Lazarus Long and the 'world' of Heinlein's novel. We find it highly implausible to suppose that

---

formulae containing free variables is "at best a convenience". He might well have added that at worst it puts us into the logical bind sketched above. For some pertinent cautionary morals about the construction and interpretation of axiomatic systems see Hughes and Cresswell, *An Introduction to Modal Logic,* 2nd ed., London, Methuen, 1968, p. 182.

75. Chapter 2, section 8. A possible-worlds parable is a story which presents a counterexample to some thesis of the form A $\longrightarrow$ B by describing a possible world in which A is true and B is false.

76. The homogeneous-worlds view, it is worth noting, is one to which quite a number of philosophers are drawn for reasons which have nothing to do with a defence of the Barcan formulae. Wittgenstein seems to have adopted it in his *Tractatus Logico-Philosophicus* (see, especially, 2.002 and 2.023). And others, like A.N. Prior, seem drawn to it because of views they hold about naming. See his *Objects of Thought*, Oxford, Clarendon Press, 1971, pp. 169–170.

everything which does exist exists necessarily and that nothing could even possibly exist except what does exist. But whichever view one adopts, this much is clear. Possible worlds — talk about which, we have argued, plays a fundamental role within propositional logic in explicating the notions of implication, validity, and the like — *continues* to play the same sort of role within predicate logic even though, at that level of analysis, the concept of a possible world itself becomes a prime object for further analysis.

That said, we turn to a question even more vexed than any we have considered hitherto, viz.,

Is there really a logic of concepts?

The case for saying that the science of logic needs to be pursued, on occasion, to a deeper level of analysis than that provided for within Predicate Logic — the Logic of Unanalyzed Concepts, as we have called it — stems from three seemingly undeniable facts: (1) that philosophers find it natural to speak of certain concepts, e.g., that of being a sister, standing in the relation of *implication* to others, e.g., that of being female; standing in this relation of *inconsistency* to others, e.g., that of being male; and so on; (2) that relations such as those of implication and inconsistency are paradigms of *logical* relations; and (3) that the analytical and notational resources of Predicate Logic (and a fortiori also those of Propositional Logic) do not suffice as ways of *justifying* our beliefs that these logical relations do in fact obtain.

Prima facie, the case is a strong one. It can be made even stronger if we turn from the now-hackneyed examples of being a sister and being female to other examples which seemingly also demonstrate the inadequacy of Predicate Logic to certify the full range of logical relations. Here is a handful of illustrative concept-pairs:

(a) *(4.45)* knowing that P
and *(4.46)* believing that P;

(b) *(4.47)* being red
and *(4.48)* being colored;

(c) *(4.49)* being an event
and *(4.50)* occurring at some time or other;

(d) *(4.51)* being taller than
and *(4.52)* being at least as tall as;

(e) *(4.53)* being more than 20 in number
and *(4.54)* being at least 19 in number.

Each of the implication relations obtaining in (a), (b), (c), (d), and (e) is representative of a whole set of similar implication relations: that in (a), of the sorts of implications which have been recognized in traditional epistemology and are nowadays enshrined in epistemic logics; that in (b), of the sorts of implications which hold between *determinate* properties and the more general *determinable* properties under which they fall; that in (c), of the sorts of implications which hold between *categories* of things and the various determinable properties which are among their essential properties; that in (d), of the sorts of implications which can hold between relational concepts; and that of (e), of the sorts of implications which can hold between quantitative and number concepts. And it would not be hard to cite examples of other pairs of concepts the members of which stand to one another in still other logical relations (drawn from the set of fifteen depicted by the worlds-diagrams of figure *(1.i)*) which fall outside the certificatory competence of Predicate Logic.

Nor can this failure be excused by saying that Predicate Logic is designed only to display logical relations between (whole) propositions rather than those between conceptual constituents of propositions. This is no excuse; it is part of the complaint. In any case, the failure is equally evident within the field of propositions. We need only consider certain propositions within which the concepts cited in (a) through (e) feature in order to see that valid inferences may be drawn from propositions, as well as concepts, in ways which Predicate Logic seemingly cannot explain. There seems no doubt, for instance, of the validity of each of the following inferences:

(a*)	from	*(4.45*)*	The Pope knows that P
	to	*(4.46*)*	The Pope believes that P;
(b*)	from	*(4.47*)*	This liquid is red
	to	*(4.48*)*	This liquid is colored;
(c*)	from	*(4.49*)*	John described the event
	to	*(4.50*)*	John described something that happened at some time or other;
(d*)	from	*(4.51*)*	Molly is taller than Judi
	to	*(4.52*)*	Molly is at least as tall as Judi;
(e*)	from	*(4.53*)*	There are more than 20 apples in the basket
	to	*(4.54*)*	There are at least 19 apples in the basket.

Yet, on the face of it, the validity of each of these inferences can be certified only by analyzing concepts which Predicate Logic must perforce leave unanalyzed.

Thus it is that the very same sorts of considerations which led us to make the move from Propositional Logic to Predicate Logic seem to impel us to make a further move from Predicate Logic, within which the logical powers of many concepts go unrecognized, to a still deeper level of logical analysis — that of a Logic of Concepts. Not surprisingly, therefore, many philosophers — especially over the past twenty years or so — have thought it wholly proper to entitle, or subtitle, their analytical inquiries "The Logic of . . . " (where the gap is filled in with a description of a concept or set of concepts, e.g., ". . . Decision", ". . . Preference", ". . . Pleasure", ". . . Religion", ". . . Moral Discourse", etc.). For instance, Jaakko Hintikka subtitles his book *Knowledge and Belief* — one of the foundational works in epistemic logic — *An Introduction to the Logic of the Two Notions,* and takes pains to insist: "The word 'logic' which occurs in the subtitle of this work is to be taken seriously."[77] He goes on to show that logical relations of consistency, inconsistency, implication, and the like, hold between various epistemic notions ("concepts" as we have called them) in ways of which formal logic takes no cognizance.

Yet there are many logicians for whom this talk of a Logic of Concepts is, at best, to be taken in jest. At worst, they would say, such talk betrays an ignorance of the true nature of logic. The science of logic, as they see it, is a purely formal one, akin to pure mathematics, and hence has nothing to do with the properties of, or relations between, such substantive concepts as those of knowledge and belief, being red and being colored, or the like. It is concerned solely with formulating the principles or rules of valid inference which warrant certain patterns or forms of argument independently of any

77. J. Hintikka, *Knowledge and Belief: An Introduction to the Logic of the Two Notions,* Ithaca, Cornell University Press, 1962, p. 3.

special attributes of the substantive concepts which feature therein. Concepts such as those of knowledge and belief may well feature within valid arguments; but not in ways which are relevant to those arguments' validity. If an argument is valid, it is valid solely by virtue of its form. Against those who, like ourselves, are convinced that the inference

(a*) from *(4.45*)* The Pope knows that P
to *(4.46*)* The Pope believes that P

as it stands, is valid, those who believe in the omnicompetence of formal logic to deal with all matters of validity would argue: (1) that this inference is *not* valid as it stands, since it is not warranted by any rules of formally valid inference; (2) that our conviction to the contrary stems from the fact that we are taking for granted the truth of the further premise

(4.55) If any person knows that P then that person believes that P;

and (3) that when this further premise is explicitly invoked, the strictly invalid inference in (a*) is transformed into an inference whose validity Predicate Logic can easily demonstrate. For then, the two premises *(4.45*)* and *(4.55)* can be seen to exhibit the forms

$$\mathbf{K}a$$

[where the letter "a" is an individual constant[78] standing for the Pope, and the predicate letter "**K**" stands for the property of being a person who knows that P] and

$$(x)(\mathbf{K}x \supset \mathbf{B}x)$$

[to be read as "For any x, if x has the property **K** then x has the property **B**"], respectively. And the validity of the inference from these two premises to the conclusion *(4.46*)* — symbolized as "**B**a" [where "a" stands for the Pope, as before, and "**B**" for the property of being a person who believes that P] — can then be demonstrated as follows:

(1)	$\mathbf{K}a$	[Premise]
(2)	$(x)(\mathbf{K}x \supset \mathbf{B}x)$	[Unstated Premise]
(3)	$\mathbf{K}a \supset \mathbf{B}a$	[(2) × U.I.]
(4)	$\mathbf{B}a$	[(3), (1) × Modus Ponens]

But within the argument, as thus laid out, the concepts of knowledge and belief — on whose internal connections the validity of the inference was initially supposed to hinge — have dropped out of sight and out of mind. The predicate letters "**K**" and "**B**" could stand for any properties whatever and the individual constant "a" for any item whatever, and the argument would still be valid, i.e., formally valid. Moreover, it would be claimed, the same sort of treatment suffices to bring *all* cases of allegedly nonformally valid inferences within the compass of formal logic.

Now it must be admitted that the formalist's stratagem does work, in the sense that it is always possible, in the case of any example that we might cite of a nonformally valid inference, to cite some further premise or premises the addition of which will transform the inference into a formally valid one. However, this does not in itself settle the issue. For the nonformalist will be quick to point out

78. The first few lowercase letters of the alphabet are standardly used as names of particular items or individuals. They are known as *individual constants* since they are taken to have constant reference to the individuals of which they are the assigned names.

that these additional premises, which the formalist claims are needed if the inference is to be validated, are not really needed at all. After all, when we take a look at these additional, allegedly needed premises, we find that they invariably have the character of so-called analytic propositions, i.e., propositions which can be certified by analysis as necessarily true. Plainly, the additional premise, viz., *(4.55)*, which is supposedly required for the formal validation of (a*), is analytic (and hence necessarily true). And so, too, are those which are supposedly required for the formal validation of (b*) through (e*), viz., respectively,

(4.56) If anything is red then it is colored;
(4.57) If anything is an event then it happens at some time or other;
(4.58) If x is taller than y, then x is at least as tall as y;

and

(4.59) If there are more than 20 items then there are at least 19.

But it is easily shown — the nonformalist continues — that necessarily true propositions can always be dispensed with (or deleted) in the case of a valid inference.[79] Hence, if — as the formalist allows — the inferences in (a*) through (e*) are valid in the presence of these necessarily true propositions, they must also be valid in their absence.

What does the formalist have to say to all this? He will not contest the claim that if a valid inference contains a necessarily true premise, then that premise may be dispensed with without affecting the validity of the inference. For this result is one whose truth he recognizes from having examined formal systems containing formally certifiable necessary propositions. But what he will contest is the claim that propositions *(4.55)* through *(4.59)* are genuine examples of necessarily true propositions. He will allow that, within the long-standing tradition founded by Kant, they are paradigm examples of analytic propositions. But he will deny that they meet the requisite conditions for saying that they are necessarily true. For, he will now insist, a proposition can no more be said to be necessarily true unless it is formally true than an argument can be said to be valid unless it is formally valid.

At this point the dispute begins to sound as though it has come full circle, or close to it. Or rather, it begins to sound as though it is bedeviled by a large measure of verbal disagreement. What one party counts as a valid inference the other does not, since it does not meet certain formal criteria of validity; what one party counts as a necessarily true proposition the other does not since, again, it does not meet certain formal criteria — this time of necessary truth. In short, what one party counts as a logical property or a logical relation, the other does not.

The disagreement, although verbal, is not trivial. It stems from the presence, within the logical tradition established by Aristotle, of two different though related strands of concern: concern, on the one hand, with the *semantic* questions as to what it is for an argument to be valid or for a proposition to be necessarily true; and concern, on the other hand, with discovering formal or *syntactic* marks, the presence of which offers assurance of an argument's validity or a proposition's necessary truth.

There can be little doubt that, throughout much of the history of logic, the second sort of concern has been predominant. To be sure, Aristotle wrestled for some time with semantic questions about the notions of validity and necessary truth. But he did not advance much beyond the point of seeing that the first can be explicated in terms of the second — that an argument is valid when its conclusion

79. The argument is a simple one. To say that the conjunction of a proposition P with an "additional" proposition R implies a proposition Q is just to say that all the possible worlds in which P and R are true together are worlds in which Q is true. But in the case where R is necessarily true, the set of possible worlds in which P and R are true together is precisely the same set of worlds in which P is true alone. (This is easily verified by considering figure *(5.d)* in chapter 5.) Hence, if P and R imply Q, and R is necessarily true, P by itself implies Q.

follows "of necessity" from its premises — and that the second can be explicated in terms of the notion of possibility — that which is necessarily true is that which is not possibly false. His greatest achievements came with the discovery of certain formal marks of validity and his formulation of formal principles or rules which can guarantee the validity of syllogistic inferences. It was these achievements which his latter-day successors, Boole, Frege, Russell, and company, followed up so brilliantly in order to establish formal logic as a science comparable in rigor, power, and abstractness to the science of mathematics. Indeed, so preoccupied have some logicians become with the development of formal systems and techniques that, in the idiom of many, talk of logic is taken to be synonymous with talk of formal logic, or even of mathematical logic.

Too strong a predilection for the formal, however, tends to obscure the fact of the continuing presence throughout the history of philosophy of the other set of concerns: concerns with the semantic analysis of our preformal intuitions about validity, necessity, and other related logical concepts. Aristotle, we have suggested, was motivated to undertake his formal inquiries just because of the light which he thought they could throw on these concepts. And medieval logicians undertook their studies of modal logic partly for the same sort of reason. But it is only recently — since the early 1960s, in fact — that the imbalance of the formal over the semantical has begun to be redressed. It is being redressed, of course, by the development — in the hands of philosopher-logicians like Saul Kripke and Jaakko Hintikka — of so-called *possible worlds semantics.*

The merits of the possible worlds approach to logic are becoming increasingly clear to philosophers and logicians alike. It makes possible a semantical explication of the concepts of validity, necessary truth, and so on, which is free of the constraints of formal logic as hitherto conceived. As we have seen, it tells us that a proposition is necessarily true if and only if it is true in all possible worlds — an explication which accords well with Aristotle's view that necessary propositions are such that it is not possible that they should be false; and it tells us that an argument is valid if and only if in all possible worlds, if any, in which its premises are true its conclusion is true — an explication which accords well, again, with Aristotle's view that an argument is valid when its conclusion follows "of necessity" from its premises. It allows, of course, that satisfying certain formal conditions is a *sufficient* condition of an argument's validity or a proposition's necessary truth. But it does not allow the formalist's claim that these formal conditions are *necessary* ones. Thus it enables us to make good sense, for instance, of talk about knowledge implying belief without resorting to the formalist's ad hocery of invoking 'additional' premises. And it enables us to make good sense of talk about the necessary truth of propositions such as *(4.55)* through *(4.59)*, despite the fact that they are neither among the recognized truths of formal logic nor even instantiations of such truths. The explications of logical concepts offered by possible worlds semantics allow room for our belief that there is, after all, a legitimate field of logical inquiry which, for want of a better description, may be called The Logic of Concepts.

It must not be thought, however, that the possible worlds approach to the science of logic turns its back on the hard-won achievements of formal, 'mathematical' logic. On the contrary; it takes the results of formal logic for granted, gives them a semantical underpinning, and tries to supplement these results with results of its own — results which allow for the development along semantical-cum-formal lines of logics for concepts such as those of knowledge and belief, preference, decision, and so on. In short, the possible worlds approach to logic — replete as it is with modal talk — brings together the two main strands of logical inquiry in such a way that justice is done both to the achievements of the formalists and to the nonformal analyses which philosophers have traditionally given of the substantive concepts which figure centrally in our thinking about this and other possible worlds.

From the vantage point of this perspective it can be seen that, although our own terminological preferences are clear, it does not really matter how one uses the word "logic" — whether in such a way that the "formal" in "formal logic" becomes redundant or in such a way as to allow the possibility of nonformal logical attributes. What matters is only that one recognizes that concepts and propositions can have properties and stand in relations which are explicable in terms of their application or truth in the set of all possible worlds, even when those properties and relations are not recognized within established formal logics. Failure to recognize this fact can only be a stumbling block in the way of future logico-philosophical inquiry.