by him—or by me. On this moderate view, (i)–(iii) would record bodily, muscular, and neural elements of my pushing my foot downwards—not subcutaneous carryings-on.

I've sketched a language game where we single out various sorts of event components in behavior. It differs from the game in which we isolate each act token of a given kind and decide whether non-synonymous definite descriptions or reports mark a given act token. That is where multipliers should contend with radical and moderate unifiers. A third, comparatively trivial game is demarcating homogeneous spatial and temporal segments of an act token. Goldman seems engaged in yet a fourth enterprise—both interesting and legitimate: cataloguing properties exemplified during some phase of an agent's career. Evidently properties are neither ingredients nor segments of action. But perhaps Goldman imagines that individuating them amounts to individuating the deeds in which they are manifested.

As for Davidson, how should a moderate unifier persuade him that our acts comprise anything more than bodily motions? Davis (IA) and Thomson (TK) have marshalled reminders against his doctrine that effects are not to be included in what we do. I throw in the metaphysico-conceptual argument from section II. Suppose objects and the events they figure in are complementary (Davidson, IE 226–227; cf. Goldman, 773). By analogy, aren't bodily motions, moving our bodies, and thereby affecting things about us, equally essential to our notions of what it is to act and to be responsible?

IRVING THALBERG

University of Illinois at Chicago Circle

ON ENGENDERING AN ILLUSION OF UNDERSTANDING

MY title echoes a phrase from Quine. Though he himself may not have set the stage for many of the still current debates on the philosophy of logic, he certainly has illuminated it with his searching criticism and his tireless demands for rigor and clarity. The complete quotation from Quine is a familiar one and provides a convenient curtain raiser for this symposium:

* To be presented in a joint APA-ASL symposium on Entailment, December 27, 1971. Co-symposiasts will be R. K. Meyer and H. P. Grice; for Professor Meyer's paper, see this JOURNAL, this issue, pp. 808–818; Professor Grice's paper is not available at this time.

An NSF grant partially supported work on this paper. Princeton University has provided funds which make it possible to publish it at its present length. The initial ideas evolved during a graduate seminar conducted at Princeton in the fall of 1970. The author wishes to thank Javier Arratía and Philip Hansen for suggestions and for help in checking various references and details.

> Lewis founded modern modal logic, but Russell provoked him to it. For whereas there is much to be said for the material conditional as a version of 'if-then', there is nothing to be said for it as a version of 'implies'; and Russell called it implication, thus apparently leaving no place open for genuine deductive connections between sentences. Lewis moved to save the connections. But his way was not, as one could have wished, to sort out Russell's confusion of 'implies' with 'if-then'. Instead, preserving that confusion, he propounded a strict conditional and called *it* implication.
>
> It is logically possible to like modal logic without confusing use and mention. You could like it because, apparently at least, you can quantify into a modal context by a quantifier outside the modal context, whereas you obviously cannot coherently quantify into a mentioned sentence from outside the mention of it. Still, man is a sense-making animal, and as such he derives little comfort from quantifying into modal contexts that he does not think he understands. On this score, confusion of use and mention seems to have more than genetic significance for modal logic. It seems to be also a sustaining force, engendering an illusion of understanding.[1]

In the ten years since these words were written the thrusts of certain of Quine's points may have slightly dulled. *Some* would claim now a coherent understanding of modal contexts in the light of the extensive work on possible-worlds semantics. And we ought to take them seriously, for they have done *something* coherent. There remains a nagging doubt in my mind, however, that the problem solved by the use of the possible-worlds approach is *not quite* the original problem of modal logic, only an analogous one sharing certain formal similarities. In a different direction there are now those who would rashly encourage us to quantify *directly* into quotation contexts with the (to me, highly dubious) assistance of *substitutional quantification*. But that scheme—and the allure it presently holds for Quine—is a subject for another symposium.

There remain Quine's pithy complaints about the use/mention, 'if-then'/'implies' confusions, which are most accurately aimed at Russell and Lewis. I have no desire to defend confusion, even from good sources; but I would like to reopen discussion on the merits of strict implication, or better on the program that Lewis envisaged. My claim is going to be that *some* program of formalizing a "logical implication" can be presented in a nonconfused manner arriving eventually at certain of Lewis's systems. The path is different from Lewis's, but the results seem quite satisfactory; and the method—at least to me—engenders a comfortable illusion. As a consequence of

[1] "Reply to Professor Marcus," in W. V. Quine, *The Ways of Paradox and Other Essays* (New York: Random House, 1966), p. 175.

the approach we will even find something good to say about the material conditional, not functioning *as* 'implies', but used *as a parallel to*.

I

Just how confused was Russell? In some respects it seems the answer does have to be: *very*. In view of his lack of sensitivity to the needs and details of formal grammar, it is hard to deny Gödel's sharp judgment[2] that the *Principia* "presents in this respect a considerable step backwards as compared with Frege." But we should recall that at the time of the writing of the *Principia* the distinction between the logical and semantical paradoxes had yet to be made, that the object language/metalanguage boundary had not yet been sharply drawn, and, above all, Gödel's startling results were some twenty years away.

Besides this historical situation, the line between use and mention is in any case a devilishly difficult line to toe. In a given sentence you can hardly see the words for the senses. Anyone who has ever tried to proofread a manuscript knows the problem all too well. A word is barely encountered before we begin to pull out a distracting tangle of connections, associations, and images. The echo chamber of the mind is far noisier than that of any electronic rock-and-roll band. But I do not mean to give excuses.

More soberly, we can point out that the slow development of algebra in Western thought helped in its own way to create the confusion between symbols and their denotations. And the advance from Boole to Whitehead was not all that great in rigor. As Quine points out in his interesting essay on Whitehead,[3] there was a very serious muddle in the treatment of the meaning of equations in the earlier work, *Universal Algebra*, which shows up again in the curious use of the symbols 'ω' and 'j' in the sections of that book on logical algebra. It seems reasonable to me that these ingrained habits would carry over to the *Principia*. It was not Russell's contribution to algebra that excited Whitehead to collaboration; rather it was surely his use of the concepts of quantification, description, propositional functions, and relations that opened up the vision of a logical reconstruction of mathematics, thereby completing the work of Frege and Peano. This seems clear from the advanced state of Russell's thought in *The Principles of Mathematics*.

[2] "Russell's Mathematical Logic," in P. A. Schilpp, ed., *The Philosophy of Bertrand Russell* (New York: Harper & Row, 1963), p. 126.

[3] "Whitehead and the Rise of Modern Logic," in Quine, *Selected Logic Papers* (New York: Random House, 1966), pp. 9 ff.

But getting back to use and mention, let us examine a typical passage from the *Principia*:

> The Implicative Function is a propositional function with two arguments p and q, and is the proposition that either not-p or q is true, that is, it is the proposition $\sim p \vee q$. Thus if p is true, $\sim p$ is false, and accordingly the only alternative left by the proposition $\sim p \vee q$ is that q is true. In other words if p and $\sim p \vee q$ are both true, then q is true. In this sense the proposition $\sim p \vee q$ will be quoted as stating that p implies q. The idea contained in this propositional function is so important that it requires a symbolism which with direct simplicity represents the proposition as connecting p and q without the intervention of $\sim p$. But "implies" as used here expresses nothing else than the connection between p and q also expressed by the disjunction "not-p or q." The symbol employed for "p implies q," i.e., for "$\sim p \vee q$," is $p \supset q$." This symbol may also be read "if p, then q." The association of implication with the use of an apparent variable produces an extension called "formal implication." This is explained later: it is an idea derivative from "implication" as here defined. When it is necessary explicitly to discriminate "implication" from "formal implication," it is called "material implication." Thus "material implication" is simply "implication" as here defined. The process of inference, which in common usage is often confused with implication, is explained immediately.[4]

Or again we have:

> The simplest example of the formation of a more complex function of propositions by the use of these four fundamental forms is furnished by "equivalence." Two propositions p and q are said to be "equivalent" when p implies q and q implies p. This relation between p and q is denoted by "$p \equiv q$." Thus "$p \equiv q$" stands for "$(p \supset q) \cdot (q \supset p)$." It is easily seen that two propositions are equivalent when, and only when, they are both true or are both false. Equivalence rises in the scale of importance when we come to "formal implication" and thus to "formal equivalence." It must not be supposed that two propositions which are equivalent are in any sense identical or even remotely concerned with the same topic (*loc. cit.*).

Talk like this is so easy; I have to admit I do it all the time, though I think I know better. Even Reichenbach has been known to talk this way.[5]

How does Lewis fare? Not much better, I am afraid. Consider this passage from *Symbolic Logic*, noting that these quotations from

[4] *Principia Mathematica* (Cambridge: University Press, reprinted 1960; hereafter called PM), vol. 1, p. 7.

[5] In his article, "Bertrand Russell's Logic," in Schilpp, *op. cit.*, pp. 23–54; I refer here to the paragraph at the bottom of p. 27.

various authors are being presented both for their form and for their content:

> This translation [from the Boole-Schröder Algebra to the Two-valued Algebra] will serve two purposes: first, it will set off the propositional interpretation of principles which are common to the two systems from their interpretation for classes. And second, it will introduce the reader to the symbolism most used to-day for the calculus of propositions. These changes of symbolism include the following:
> (1) Instead of a, b, c, etc., we shall use p, q ,r, etc., for propositions.
> (2) Instead of $-p$, we shall write $\sim p$ for the contradictory of p, or "p is false."
> (3) Instead of $p \subset q$, we shall write $p \supset q$ for "If p is true, then q is true" or "p implies q."
> (4) Instead of $p + q$, we shall write $p \vee q$ for "At least one of the two, p and q, is true" or "Either p or q."
> (5) Instead of $p = q$, we shall write $p \equiv q$ for "p and q are both true or both false" or "p is equivalent to q." However, we shall continue to use the symbol $=$ for purely notational equivalences, such as $p = 1$, and for indicating that two expressions are to have the same meaning, by definition.[6]

The translation from *inclusion* to *implication* is misleading enough, but the remarks about equality are just incoherent. Scarcely more than a page later we find:

> A proposition p is either always true or always false; either $p = 1$ or $p = 0$. The symbolism $\sim p$ means "p is false"; so does $p = 0$. Hence these two are here interchangeable. Similarly p and $p = 1$ are different ways of expressing "p is true," and are interchangeable. We can, then, wherever we choose, write $\sim p$ for $p = 0$, in a theorem, and p for $p = 1$; the meaning of the theorem will be unaffected. Thus the expressions $p = 0$ and $p = 1$ can be eliminated. However, we shall sometimes write a principle in both ways, in order that the derivation of the simpler form may be clear (82).

Have I missed the point? I see nothing to do with "notational equivalences" here.

In the *Principia* Russell and Whitehead proceeded *from* the propositional calculus *to* class algebra; Lewis and Langford in *Symbolic Logic* returned to the earlier sequence of presentation, I suppose for pedagogical and historical reasons. There is something to be said, however, for the *explanation* of class algebra through propositional functions, whereby a separation between the *relation* of inclusion and

[6] C. I. Lewis and C. H. Langford, *Symbolic Logic* (New York: Century, 1932; Dover reprint, 1959, hereafter called SL), p. 80.

a certain Boolean *operation* (defined by the polynomial '$- p + q$') is quite necessary. But when the Boolean algebra *is* the algebra of the two truth values, then the distinction vanishes, since now every Boolean operation is in fact a relation. Even *equivalence* and *equality* are the same for the truth-value algebra. Maybe this is where part of the confusion sets in; especially if no distinction is drawn between the two-element class algebra and the two-element truth-value algebra.

The other side of the difficulty, however, revolves about the ideas of *conditional*, *implication*, and *inference*. Norbert Wiener tried once to defend Russell against Lewis on the basis of Russell's notion of *formal implication*[7] (a universally quantified material conditional). There is a long passage in the *Principia* on this idea which Russell considered very important (20f.). It looks like quicksand to me, however, particularly since Russell brings *knowledge* into the discussion early on. In any case, in a conditional such as:

$$\varphi(x,y) . \supset . \psi(x,y)$$

how are we supposed to tell whether quantification on 'x', or on 'y', or on both makes this "formal"? Russell is confounding the propositional function with its value—in the way so common to mathematicians with their use of such expressions as '$f(x)$'—as well as mixing up use and mention. It does not seem to me that there is any hope for the idea of formal implication without bringing in consideration of the metalanguage and properties of the formal provability of sentences. But it is amusing to note that, from the possible-worlds point of view, universally quantified conditionals would behave just like strict implications in the system **S5** of Lewis.

How did Lewis come to favor strict implication? He first listed all the peculiar properties of the material conditional, and we do not need to review here these so-called "paradoxes" today. Then he began the elimination of the undesirable tautologies with the aid of the concept of an *unanalyzed proposition* leading to the derivation of "a calculus of propositions from symbolic postulates, by the logistic method" (SL, 122). You see, he wanted 'p implies q' to mean 'q is deducible from p', but he wanted to use 'p' and 'q' in the familiar way as propositional variables. He probably did not even know about meta- or sentence variables. (And he did not seem to want to learn about them either, to judge from Quine's remarks.) Since his propositions were to be unanalyzed, his calculus would "comprise only the most general principles of deductive inference" (*loc. cit.*) That is, the

[7] "Mr. Lewis and Implication," this JOURNAL, XIII, 24 (Nov. 13, 1916): 656–662.

results would "not depend upon the analysis of the propositions in question" (*loc. cit.*), in the way that the highly extensional view whereby the value of a proposition is a truth value no doubt leads to.

Strict implication itself is defined in terms of *possibility*:

11.02 $$p \prec q . = . \sim \Diamond (p \sim q)$$

but all the *axioms* are given in terms of '$\prec$'. Lewis says:

> The properties of this relation, and its precise significance, will become clearer with the development of the system, particularly in Section 4 (SL, 124).

I pass over what he says about *logical equivalence*; this is the "logistic method" indeed! I do not mean to be too unkind, however; the system with its axioms, rules of inference, and sequence of theorems is very nicely set out, and it *is* fun to do. Come to think of it, though, that may just be one of the dangers of the formal approach to thought: it is too easy to be tidy and elegant without having sufficient justification for the particular systematization.

When we look at the promised Section 4, we find at last a discussion of the modal function denoted by '$\Diamond p$' and read "p is possible" or "p is self-consistent." And after a long sequence of theorems strict implication is found to be equivalent to the *necessary* material conditional (*18.7*). But is anything clearer now besides these formal dependencies? Lewis agrees that "the words 'possible,' 'impossible', and 'necessary' are highly ambiguous in ordinary discourse" (SL, 160). But he continues, "The meaning here assigned to $\Diamond p$ is a *wide* meaning of 'possibility'—namely, logical conceivability or the absence of self-contradiction" (160/1). Such colloquial renderings are often helpful in following the formal development, but I cannot see how such remarks can serve in place of a serious justification. In this circumstance we are led to ask whether there could be *any* justification at all for such a confused concept as strict implication.

II

It was very reasonable of Lewis to ask for the most general principles of deductive inference, even if we come to judge him as not having supplied them himself. Fortunately a very good answer has been found subsequently: namely, by Gentzen. *Un*fortunately, in my opinion, both because of the aims of Gentzen's own work and in the light of later applications, the Gentzen systems have been very much oriented toward proof-theoretic analyses—especially the problems of establishing the so-called *cut-elimination theorem*. For me this was misleading. It took me a long time to realize that cut is *not* eliminable—except in very special circumstances. This is not to say that cut elimination is uninteresting or unimportant, but there does seem to be a simple and basic point to make with the aid of Gentzen's idea which may not be so generally appreciated.

Let us side with Lewis in asking for a calculus of unanalyzed propositions, but let us at the same time try to be clearer about what we are doing. Maybe the word 'proposition' is worn out from long misuse. Can we find a better one? I have no brilliant suggestion, but I would like to propose 'statement'. For reasons that will become apparent in a moment, we shall see that 'sentence' will not do. I want instead to use 'statement' in a way that is *intermediate* between 'proposition' and 'sentence'. The latter is too syntactical; the former carries around too heavy a load of metaphysical baggage. The choice of 'statement' is not perfect, either, for it would seem to harbor within its meaning some sense of the *action* of stating something, a sense I do not intend. Possibly 'potential statement' would be closer to my intention, but that phrase is clumsy. So let me request your good will for the moment in using 'statement' in a rather neutral manner.

For the sake of systematization, let us suppose that we are considering a "language" for which the relevant statements form a *set* $\mathcal{S}$. Now the elements $A \in \mathcal{S}$ are but potential statements, and something active has to be done to assert one of them. This is embodied in the notion of *assertion* symbolized by the sign '$\vdash$'. Thus, to say

$$\vdash A$$

is to assert the *truth* of A, usage taken over from Frege and popularized by the *Principia*. As everyone knows, Gentzen, building on earlier work of Hertz, extended the notion to that of *conditional assertion*. Thus we can say:

$$A \vdash B$$

or

$$A_0, A_1, \ldots, A_{n-1} \vdash B$$

or even

$$A_0, A_1, \ldots, A_{n-1} \vdash B_0, B_1, \ldots, B_{m-1}$$

in the style of Gentzen's well-known calculus of sequents.[8]

Are we making any sense? Yes, having first specified the set $\mathcal{S}$ of statements, we next specify $\vdash$. What is this $\vdash$? It is a *relation*, the relation of conditional assertion; specifically it is a binary relation between *finite sets* of statements. This last convention is but a con-

[8] Cf. *The Collected Papers of Gerhard Gentzen*, W. E. Szabo, ed. (Amsterdam: North-Holland, 1969).

venience to make the listing of the properties of $\vdash$ particularly easy. Thus when we write:

$$A_0, A_1, \ldots, A_{n-1} \vdash B_0, B_1, \ldots, B_{m-1}$$

we really mean to write:

$$\{A_0, A_1, \ldots, A_{n-1}\} \vdash \{B_0, B_1, \ldots, B_{m-1}\}$$

as a relationship between sets. If we like, we could insist that $\vdash$ is a relation between finite *sequences* of statements; but that refinement is irrelevant for the present discussion.

What could this relation $\vdash$ be? What relation is it? The answer depends on many factors: it depends on the nature of the elements of $\mathcal{S}$; it depends on what we want to do with them. But Lewis did not want us to analyze the elements of $\mathcal{S}$, and we shall not. We do, however, allow ourselves to say that statements can be either *true* or *false*, and in various ways. That does not imply any analysis, but it is part of our understanding of what it is that "propositions" convey. They may also convey other information, and we can conceive of other "languages" where no truth value at all is "embedded" in a proposition. Such considerations will lead to other relations, however; relations different from the conditional assertion of classical logic. Just how different, though, is a question that must be investigated; in certain instances the difference is not very extreme—at least for many essential properties.

Let us then agree that truth values are important and that truth values can be assigned to statements—perhaps in various ways. To make this idea precise, we can introduce the concept of a *valuation* as a function *from* statements *to* truth values. Let 1 be the truth value *true* and let 0 be the truth value *false*. Mathematically speaking, then, a valuation is a function

$$v : \mathcal{S} \rightarrow \{0,1\}$$

For each $A \in \mathcal{S}$, we have $v(A)$ as the truth value of A as assigned to A by the function v. Different ways of making assignments correspond to different functions. One way implies that each statement has a *unique* truth value; hence, our use of functions. All of this is terribly familiar, and it was meant to be. I am only trying to emphasize the proper level of abstraction to be able to meaningfully discuss Lewis's program.

What then about $\vdash$? How does this relation involve valuations? In the first place, depending on our intentions, not all valuations are reasonable. Though it may not always be immediately apparent what they are, we can postulate a set $\mathcal{V}$ of *intended valuations* appropriate to our understanding of the statements in $\mathcal{S}$. (Careful! Understanding does not imply knowledge. Let us keep knowledge out of this discussion.) It is in terms of $\mathcal{V}$ that $\vdash$ becomes completely definite, for we can now define for two finite subsets $\mathcal{A}, \mathcal{B} \subseteq \mathcal{S}$:

$$\mathcal{A} \vdash \mathcal{B}$$

to mean that every valuation $v \in \mathcal{V}$ making *all* statements in $\mathcal{A}$ *true*, makes *at least one* statement in $\mathcal{B}$ *true* also. That is, under the *condition* of truth for all $A \in \mathcal{A}$, the truth of at least one $B \in \mathcal{B}$ *follows*. It is only a conditional assertion. Note that the '$\mathcal{B}$'-part is taken as a *set*, to allow for full true-false *duality*. For we can also say that $\mathcal{A} \vdash \mathcal{B}$ if and only if every valuation in $\mathcal{V}$ making all statements in $\mathcal{B}$ *false* makes at least one statement in $\mathcal{A}$ *false* also.

Having made more precise the import of $\vdash$, what of its properties? They are easy to state and easy to appreciate:

(Reflexivity) $\qquad \mathcal{A} \vdash \mathcal{B}$ if $\mathcal{A} \cap \mathcal{B} \neq \phi$

(Monotonicity) $\qquad \dfrac{\mathcal{A} \vdash \mathcal{B}}{\mathcal{A}, \mathcal{A}' \vdash \mathcal{B}, \mathcal{B}'}$

(Transitivity) $\qquad \dfrac{\mathcal{A} \vdash B, \mathcal{C} \qquad \mathcal{A}, B \vdash \mathcal{C}}{\mathcal{A} \vdash \mathcal{C}}$

In the above, the script letters range over finite subsets of $\mathcal{S}$ and the italic capitals over the elements of $\mathcal{S}$. I have written '$\mathcal{A}, \mathcal{A}'$' as short for '$\mathcal{A} \cup \mathcal{A}'$', and '$\mathcal{A}, B$' as short for '$\mathcal{A} \cup \{B\}$'.

We can call these laws '(*R*)', '(M)', and '(T)' for short. What is the status of these laws? They are all statements of the *metalanguage*, counting the statements in $\mathcal{S}$ as being those of the *object language*. Besides this, granting the definition of $\vdash$ in terms of $\mathcal{V}$, they are all *true*. I have written (M) and (T) as *inference rules*, but that is just a shorthand for the material conditional from the conjunction of the premises to the conclusion. A person does not have to know about the *theory* of material conditionals to be able—in the metalanguage—to apply these two specific conditionals in an argument concerning the relation $\vdash$. Rule (T) is the famous *cut rule*, and, far from being mysterious, it is just a very characteristic and valid property of conditional assertion.

Are there other valid rules? Yes, of course. For example:

$$\frac{A \vdash B \qquad B \vdash C \qquad C \vdash D}{A \vdash D, E}$$

is clearly valid; for if A is made *true*, then in turn so are B, C, and D. But obviously it can be *proved* on the basis of (R), (M), and (T). We are asking the *metametaquestion* whether there are any properties of a relation of conditional assertion that are not implied by (R), (M), and (T). In other words, the question is whether these three laws provide an adequate axiomatization in the metalanguage of the simple properties of conditional assertion. If so, we then have what seems to me to be a solution to a big part of Lewis's program of setting up—as a metalinguistic activity—a general theory of deductive inference.

The essentials of the argument for the completeness of the rules are well known; I give only the briefest summary. In a completely abstract setting we suppose we have a set $\mathcal{S}$ and a relation $\vdash$ between finite subsets of $\mathcal{S}$ which is assumed to satisfy (R), (M), and (T). We ask: could such a relation $\vdash$ have come from a set of valuations? The answer is *yes*, and we can even find the *largest* such set of valuations. Given $\vdash$, we say that a valuation v is *consistent with* $\vdash$ (in symbols: $v \in \mathcal{V}_{\vdash}$) if and only if whenever $\mathcal{A} \vdash \mathcal{B}$ and v makes all statements in $\mathcal{A}$ *true*, it makes some statement in $\mathcal{B}$ *true* also. One then proves along standard lines that the given relation $\vdash$ *could* have been defined in terms of $\mathcal{V}_{\vdash}$ as the set of intended valuations. One has to prove that if $\mathcal{A} \vdash \mathcal{B}$ does not hold, then there is a valuation $v \in \mathcal{V}_{\vdash}$ making all of $\mathcal{A}$ *true* and all of $\mathcal{B}$ *false*. In case the set $\mathcal{S}$ is *countable*, the argument is a simple stepwise construction of v. In case $\mathcal{S}$ is *uncountable* (if one allows such abstractions), then the proof requires a "maximum" argument established with the aid of the axiom of choice. When I said before that cut cannot be "eliminated", I meant that in this proof we cannot leave off the assumption that $\vdash$ satisfies (T).

Is such a level of abstraction and sophistication really necessary for discussing such simple notions? Yes, I assert, it is—if one wants a general theory. And we do want the generality. For one reason we are enabled to argue about the connection between $\vdash$ and $\mathcal{S}$ once and for all. The point is that it covers all the cases. Thus the set $\mathcal{S}$ could be construed in one instance as comprising the *sentences* of a formal language, so that the elements of $\mathcal{S}$ are *syntactical* objects. On the other hand, the elements of $\mathcal{S}$ could be *classes*, where the intended valuations correspond to points of the universal class, and where:

$$A_0, A_1, \ldots, A_{n-1} \vdash B_0, B_1, \ldots, B_{m-1}$$

is equivalent to:

$$A_0 \cap A_1 \cap \cdots \cap A_{n-1} \subseteq B_0 \cup B_1 \cup \cdots \cup B_{m-1}$$

Or we could imagine that $\mathcal{S} = \{0,1\}$, and there is only *one* intended valuation: the identity function. Or...or...or.... There are a multitude of interpretations, all encompassed within the general theory of conditional assertion.

What of the notions of *proof* and *deduction*—in the object language? Yes, they are included in the range of the theory also. Thus, given a set of *basic* (or: *axiomatic*) conditional assertions

$$\mathcal{A}_i \vdash_0 \mathcal{B}_i$$

where $i \in I$ an index set, then there is a *least* relation $\vdash$ which contains these pairs $(\mathcal{A}_i,\mathcal{B}_i)$ and which satisfies (R), (M), and (T). The original $\vdash_0$ assertions generate the $\vdash$ assertions with the aid of "formal proofs." This is the most common way we get $\vdash$ relations in formal logic—and after we obtain them we then begin asking about appropriate valuations. And here is one point at which cut elimination enters meaningfully: the generation of $\vdash$ from $\vdash_0$ may very well be possible *without* recourse to rule (T). In the end (T) is satisfied, but we did not have to "build it in" to the definition of $\vdash$. In this way some inductive proofs about $\vdash$ may be easy to carry out, or there may be other causes for rejoicing, when this elimination is possible. But, it is not always possible; and in any case no matter how $\vdash$ is defined, it must satisfy (T) to count as a relation of conditional assertion.

Another word of caution: though for certain $\vdash$ it may be appropriate to read '$A \vdash B$' as 'A implies B', 'A entails B', 'B is a consequence of A', 'B is deducible from A', or the like, it is not required to do so. In this abstract setting we have founded bivalent logic on the concepts of *valuation* and *conditional assertion*, but have left open the specification of the nature of the "statements." The theory of "inference" then applies to *all* the standard interpretations (class algebra, Boolean algebra, two-valued algebra, formal systems) as well as to many nonstandard cases including the modal languages which Lewis liked so well. We have to see, however, whether this path leads to the proper understanding.

III

Before tackling modal logic and strict implication, it might be well to review quickly how ordinary propositional calculus works. In terms of rules of inference we have:

$$(\bot)\quad \frac{\mathcal{A} \vdash \mathcal{B}}{\mathcal{A} \vdash \bot,\mathcal{B}} \qquad\qquad (\top)\quad \frac{\mathcal{A} \vdash \mathcal{B}}{\mathcal{A}, \top \vdash \mathcal{B}}$$

$$(\vee)\quad \frac{\mathcal{A} \vdash A,B,\mathcal{B}}{\mathcal{A} \vdash A \vee B,\mathcal{B}} \qquad\qquad (\wedge)\quad \frac{\mathcal{A},A,B \vdash \mathcal{B}}{\mathcal{A},A \wedge B \vdash \mathcal{B}}$$

$$(\sim)\quad \frac{\mathcal{A} \vdash A,\mathcal{B}}{\mathcal{A}, \sim A \vdash \mathcal{B}} \qquad\qquad (\supset)\quad \frac{\mathcal{A},A \vdash B,\mathcal{B}}{\mathcal{A} \vdash A \supset B,\mathcal{B}}$$

Each of the rules is *double-edged*: the assertions above and below the double line are considered equivalent, as inferences the rules operate both ways.

An immediate question upon seeing the formulation of these rules is whether I am making a gaffe in use-mention etiquette. Quine's first reaction, I am sure, would be to say 'yes', but I can defend myself. When Quine writes a rule with quasiquotes:

$$\frac{\mathfrak{A}, A \vdash B, \mathfrak{B}}{\mathfrak{A} \vdash \ulcorner(A \supset B)\urcorner, \mathfrak{B}}$$

he is thinking of $\mathcal{S}$ as consisting of *syntactical objects*. The metalinguistic expression '$\ulcorner(A \supset B)\urcorner$' is his very convenient shorthand for the syntactical *operation* of filling in the blanks of '($\supset$)' with A and B respectively. We have of course assumed that $\mathcal{S}$, the set of sentences or well-formed formulas of our formal language, is closed under this syntactical operation. My defense is that I am not thinking of $\mathcal{S}$ syntactically at all. I am leaving the interpretation of $\mathcal{S}$ open.

In *abstract* propositional calculus we do not care what the "statements" in $\mathcal{S}$ actually are. We do assume that $\mathcal{S}$ contains special "constant" statements ($\perp$ and $\top$) and that $\mathcal{S}$ is closed under three binary and one unary operations ($\vee$, $\wedge$, $\supset$, and $\sim$). Besides the structure these operations give to $\mathcal{S}$, we assume further that the relation $\vdash$ satisfies all the nine rules (R), (M), (T), ($\perp$), ($\top$), ($\vee$), ($\wedge$), ($\sim$), and ($\supset$). The symbols like '$\vee$' and '$\supset$' are symbols of the metalanguage *not* of the object language. Thus if the object language had a Polish accent and were syntactical, then we might find that:

$$(A \supset B) = \ulcorner CAB \urcorner$$

for all $A,B \in \mathcal{S}$. But if the object language were class-conscious, we might find that:

$$(A \vee B) = (A \cup B)$$

Thus, by thinking metalinguistically, we see no confusion was either intended or committed.

Though the rules in the above forms are very convenient to state, one should note that in the case of conjunction, say, the rule ($\wedge$) is equivalent to the combination of these three conditional assertions:

$$A,B \vdash A \vee B$$
$$A \wedge B \vdash A$$
$$A \wedge B \vdash B$$

It then becomes obvious that no matter what $\mathcal{S}$ is, as long as the rules are satisfied,

$$v(A \wedge B) = 1 \text{ iff } v(A) = 1 \text{ and } v(B) = 1$$

holds for all $A,B \in \mathcal{S}$ and all consistent valuations $v \in \mathcal{V}_{\vdash}$. We thus conclude that rule ($\wedge$) is exactly what we need to interpret *conjunc-*

tions as *and* under consistent valuations. We can therefore argue for the completeness of the above rules in the usual way. The same fact about consistent valuations also leads directly to the *representation theorem for Boolean algebra* as another application of the method. These connections are well known, but it is worth while to see how they all fit together.

For future reference, the corresponding three assertions for the material conditional are the following:

$$A, A \supset B \ \vdash B$$
$$B \ \vdash A \supset B$$
$$\vdash A, A \supset B$$

where in the last the left-hand side is empty. It means that every consistent valuation must either make A or $A \supset B$ *true* no matter how "odd" the $A \supset B$ combination seems.

It is also important to note that the rules treat the connectives *independently*: we could assume rule ($\wedge$) and closure under $\wedge$, *without* any other conditions on $\mathcal{S}$, and still be able to make the same remarks about valuations of conjunctions. Merely because you have one connective, you are not committed to the rest. I feel this is an enormous advantage to the Gentzen method—especially when one has struggled with the tangle of dependencies among primitives, as remarked previously for the original Lewis system.

IV

I make the guess that Lewis would not be too happy with the course of this discussion up to this point. The relation $\vdash$ would make good sense to him, but I doubt he would be satisfied with our theory of inference. The trouble would be that the metalanguage remains *unformalized*; somehow the inferences take place in thin air, and he wanted to *see* them, to have them embodied in a formal calculus. There is nothing to stop us from formalizing the metalanguage. Then by a *meta*metalinguistic definition we could investigate *derived* rules of inference. (The proper definition is not immediately apparent, but that is another story.) But that was not Lewis's approach: he wanted, in effect, to inject a part of the metalanguage into the object language with the aid of his modal operators. Or at least that is my "rational" reconstruction of his plan. Quine calls it a use-mention confusion; he may be right, but let us proceed further.

We assume, at least, a set $\mathcal{S}$ closed under $\bot$, $\top$, $\vee$, $\wedge$, together with a relation $\vdash$ satisfying the appropriate rules of the previous sections. (For a gain in generality, we omit negation at the start.) The $\vdash$ relation allows inferences, but we want to see them in black and white somehow reflected in the statements of $\mathcal{S}$ itself. To this end we postulate that $\mathcal{S}$ is closed under a new binary operation $\Rightarrow$. (It seems better not to prejudice matters by using Lewis's hook.) The problem, then, is to imagine which rules of inference are appropriate for the intended role of $\Rightarrow$.

One basic rule is easy to guess; I call it '(C)':

$$\text{(C)} \qquad \frac{A \vdash B}{\vdash A \Rightarrow B}$$

The 'C' may stand either for 'conditionalization' or for 'confusion'. The idea is that if $\Rightarrow$ is going to function as a "logical" implication, then the logical validity of a conditional assertion can equivalently be "confused" with the assertion of the implicational statement.

Formally speaking, the rule (C) is a modification of rule ($\supset$) given in III. In one direction it is a weak form of the so-called *deduction theorem*; in the other direction it is a weakening of *modus ponens*. It seems better at this stage *not* to try to assume *modus ponens* in the stronger form:

$$\text{(MP)} \qquad A, A \Rightarrow B \vdash B$$

Stated in this way there is an uncomfortable mixing of levels which I prefer to avoid until I have a better grasp of the nature of the statement $A \Rightarrow B$.

As our avowed purpose in introducing $\Rightarrow$ was to make inferences "visible," we should ask what this could mean. Suppose that a certain figure is "valid" in some good sense. The shape of an instance of the figure would be, say:

$$\begin{array}{c} \mathcal{A}_0 \vdash \mathcal{B}_0 \\ \mathcal{A}_1 \vdash \mathcal{B}_1 \\ \cdots \\ \mathcal{A}_{n-1} \vdash \mathcal{B}_{n-1} \\ \hline \mathcal{A}_n \vdash \mathcal{B}_n \end{array}$$

where the $\mathcal{A}_i$, $\mathcal{B}_j$ are finite subsets of $\mathcal{S}$. By assumption of the rules ($\perp$), ($\top$), ($\vee$), ($\wedge$), the various sets can be combined into *single* statements; thus, the figure could be written as:

$$\text{[V]} \qquad \begin{array}{c} A_0 \vdash B_0 \\ A_1 \vdash B_1 \\ \cdots \\ A_{n-1} \vdash B_{n-1} \\ \hline A_n \vdash B_n \end{array}$$

where the A_i, $B_j \in \mathcal{S}$. The sense in which this inference is "valid" might very well be that it is *derivable* from "the rules." The only trouble with this approach is that we do not yet know exactly what the rules are. But, in an experimental spirit, let us suppose that the rules have already been constructed to solve the problem, and on this assumption we shall try to discover what they are.

Now if the inference [V] is valid, we want to be able to transcribe it into explicit statements in $\mathcal{S}$. The way is quite obvious:

[H] $$A_0 \Rightarrow B_0, A_1 \Rightarrow B_1, \ldots, A_{n-1} \Rightarrow B_{n-1} \vdash A_n \Rightarrow B_n$$

From the "vertical" inference [V], we have come down to earth at the "horizontal" conditional assertion [H]. In other words, [H] is the "visible" version of [V]. The general rule (I) about inferences is therefore that, from any valid [V], we may pass to the validity of the corresponding [H]. Note that the rule (C) already includes the converse; for, if an [H] is valid, the corresponding [V] may be derived by several applications of (C).

The description of (I) just given is not very satisfactory. As stated it is at least on the metametalevel and is thus of a different character from the other rules. Nevertheless, the rule is easy to use, and we give some typical examples.

In the first place, with the aid of (C) it is seen quickly that:

$$\begin{array}{c} A \vdash B \\ A \Rightarrow B \vdash C \Rightarrow D \\ \hline C \vdash D \end{array}$$

is valid. By (I) we can thus derive:

(MP$_{\Rightarrow}$) $$A \Rightarrow B, (A \Rightarrow B) \Rightarrow (C \Rightarrow D) \vdash C \Rightarrow D$$

This is a weak version of *modus ponens*: namely, a restriction to statements of the form $X \Rightarrow Y$.

As a second example, note the obvious validity of:

$$\begin{array}{c} A \vdash B \\ B \vdash C \\ C \vdash D \\ \hline A \vdash D \end{array}$$

Thus by (I) we have:

$$A \Rightarrow B, B \Rightarrow C, C \Rightarrow D \vdash A \Rightarrow D$$

But then, using (C), we derive:

$$\begin{array}{c} A \vdash B \\ C \vdash D \\ \hline B \Rightarrow C \vdash A \Rightarrow D \end{array}$$

Another application of (I) provides:

(T$_{\Rightarrow}$) $$A \Rightarrow B, C \Rightarrow D \vdash (B \Rightarrow C) \Rightarrow (A \Rightarrow D)$$

This conditional assertion is a type of *transitivity* for $\Rightarrow$.

As our third example, recall the rule (T) itself, an instance of which is:

$$\frac{\begin{array}{c} A \vdash B, C \\ A, B \vdash C \end{array}}{A \vdash C}$$

We may rewrite it as:

$$\frac{\begin{array}{c} A \vdash B \vee C \\ A \wedge B \vdash C \end{array}}{A \vdash C}$$

and then by (I) pass to:

$(\mathrm{T}_{\vee\wedge})$ $$A \Rightarrow (B \vee C), (A \wedge B) \Rightarrow C \vdash A \Rightarrow C$$

This last is a quite direct horizontalization of (T).

These examples show how applications of the rule (I) in collaboration with rule (C) generate explicit and mildly intricate conditional assertions involving the operation $\Rightarrow$. Obviously, infinitely many additional examples of the same kind may be obtained in a similar way. Fortunately, it is not necessary to worry about further examples, for these three schemata suffice for the reduction of the rule (I) down to the ordinary metalevel. More specifically, the metatheoretic result I have obtained may be formulated as follows:

> Suppose a system is set up with $\mathcal{S}$ assumed closed under $\bot$, $\top$, $\vee$, $\wedge$, and $\Rightarrow$, and possibly other operations. Suppose that the $\vdash$ relation is constrained at least by the rules:
>
> (R), (M), (T),
> ($\bot$), ($\top$), ($\vee$), ($\wedge$),
> (C), (MP$\Rightarrow$), (T$\Rightarrow$), ($\mathrm{T}_{\vee\wedge}$),
>
> Then, no matter what additional rules of the normal kind are adjoined, the system will also satisfy (I) with respect to all derived rules [V].

Strictly speaking, this result is a *metametatheorem*, because it concerns *types* of systems rather than a particular $\mathcal{S}$ and particular $\vdash$. I think its import is clear, however. The rules (MP$\Rightarrow$), (T$\Rightarrow$), (T$\vee\wedge$) were first shown to be necessary, and then they proved to be sufficient.[9]

The sense of 'sufficient' must be carefully described. A logistic system is a language $\mathcal{S}$ (with its various closure conditions) together

[9] A system somewhat similar to the one presented here can be found in W. and M. Kneale, *The Development of Logic* (Oxford: University Press, 1962), pp. 559–568. Kneale, however, confuses $\vdash$ and $\Rightarrow$ and treats $\Rightarrow$ as a multiary connective. He uses a natural-deduction formalization of his rules. He assumes (MP) and favors a strong rule that amounts to **S5**. He did not seem to be aware of the metatheorem about (I) that is the main point of our discussion.

with a relation $\vdash$ of conditional assertion (sometimes called an *entailment* relation or a *consequence* relation—depending on whether the relation or its converse is emphasized more). What is confusing is that we are seldom satisfied with the contemplation of *one* specific system; rather we axiomatize whole *classes* of systems. What is doubly confusing is that certain languages constructed syntactically (so that their statements contain "variables" or schematic letters) can be pressed into service along with suitable entailment relations to do duty as *generic representatives* of their class. (Mathematicians often speak of "free" algebras in a similar connection.) My present position is that the single-minded pursuit of these generic examples is a mistake: the interest really lies in the classes. We must, therefore, ascend to the higher language level. We are much better off thinking about what I earlier called *abstract* propositional calculus—for reasons given and, I hope, illustrated. That metatheory can be (fragmentarily) self-applied is interesting and important, but the device must be viewed in the proper setting.

For the type of $\Rightarrow$ system we have been discussing, a confusion of levels would bring instant chaos. A horizontalization from [V] to [H] makes intuitive sense only if [V] has general, not accidental validity. Hence the rule [V], as an inference, must be *derivable* from the other assumed rules—if one is going to play safe about generality. Rules are statements in the metalanguage; derivability of rules is defined in the metametalanguage. This situation points up the reason why it is better to call a metastatement:

$$\mathcal{A} \vdash \mathcal{B}$$

a *conditional assertion* rather than an *entailment relationship*. The, words 'entailment' and 'consequence' harbor an echo of syntactic shuffling—especially to us who have spent so much time fooling with formal systems. But no following of "rules" in getting from $\mathcal{A}$ to $\mathcal{B}$ ought to be intended.

What is confusing can be illustrated by a simple conditional assertion:

$$A,B \vdash C$$

If we assume this to hold in a given system, then the following inference is correct:

$$\frac{\vdash A \quad \vdash B}{\vdash C}$$

If we think of the conditional assertion as an inference without premises, then the passage is a derivation of one rule from another (with the aid of the always assumed rule, (T), of course). However, a moment's thought will show that the two rules are not equivalent: the conditional assertion is the stronger. Nevertheless, we often reach desirable conclusions by tossing around conditional assertions as if they were nothing more than the weaker rules. Success seldom requires analysis—failure always ought to.

V

We can now say something good about material implication, since the reader will have noticed that the rule ($\supset$) gives more than (C), ($MP_{\Rightarrow}$), ($T_{\Rightarrow}$), and ($T_{\vee\wedge}$). To be precise, consider any systematization based on at least the rules for classical propositional calculus. Then the general result of the last section assures us that, say, a rule of the form:

$$\frac{A \vdash B}{C \vdash D}$$

is deducible if and only if the conditional assertion:

$$A \supset B \vdash C \supset D$$

is deducible. In other words, *within certain contexts* material implication is quite adequate for making inferences "visible." It seems to me that this explains why there is generally very little need for going beyond material implication. Even though the statement $A \supset B$ does not mean in itself that A entails B, a relationship such as that displayed above between two $\supset$ statements has the same force as an inference from one entailment to another.

It seems safe to say that Lewis would not have been really satisfied by the remark just made about $\supset$. In discussing implication and deducibility he says:

> Thus not only may a relation of truth-implication give rise to inference, but also it gives rise to inference in such a manner that when the premise is a law of the system, the consequent also will be a law of the system. The principles of Strict Implication express the facts about any such deduction in an explicit manner in which they cannot be expressed within the truth-value system itself, for the reason that, in Strict Implication, what is tautological is distinguishable from what is merely true, whereas this difference does not ordinarily appear in the symbols of a truth-value system.
>
> In the light of all these facts, it appears that the relation of strict implication expresses precisely that relation which holds when valid deduction is possible, and fails to hold when valid deduction is not possible. In that sense, the system of Strict Implication may be said to provide that canon and critique of deductive inference which is the desideratum of logical investigation. All the facts about those

cases in which any other implication-relation may genuinely give rise to inference are incorporated in and explained by the laws of this system—or they may be so incorporated by introducing the implication-relation in question into the system by definition (SL, 247).

Part of the trouble in trying to accommodate all of Lewis's demands is the real confusion between *validity* and *deducibility*, which surely rests on a use/mention confusion that can be resolved only by separating the language levels. A certain amount of help can be provided with the aid of valuations and an extension of language.

Assume that a particular system with an $\mathcal{S}$ and a $\vdash$ has been given, satisfying all the rules of classical propositional calculus. We can also assume a set $\mathcal{V}$ of intended valuations as explained in section II. Thus $A \vdash B$ holds just in case it is valid in the sense that, for all $v \in \mathcal{V}$, if $v(A) = 1$, then $v(B) = 1$. Now extend $\mathcal{S}$ to a *larger* set $\mathcal{S}^*$ of statements closed under $\perp$, $\top$, $\vee$, $\wedge$, $\sim$, $\supset$, and a new operation $\Rightarrow$. (This can be done in many ways; we should of course take care that the old meanings of the classical connectives are consistently extended to the larger set.) If we are clever in how we construct $\mathcal{S}^*$, we shall be able to extend each $v \in \mathcal{V}$ to a v^* defined on $\mathcal{S}^*$ satisfying the usual conditions with respect to the classical connectives. In addition, we shall be free to specify for $v \in \mathcal{V}$ and for all $A,B \in \mathcal{S}^*$ that

$$v^*(A \Rightarrow B) = 1 \text{ iff, for all } w \in \mathcal{V}, \text{ if } w^*(A) = 1, \text{ then } w^*(B) = 1$$

Such a construction can certainly be carried out along standard lines. We thus obtain

$$\mathcal{V}^* = \{v^* : v \in \mathcal{V}\}$$

as the set of intended valuations for $\mathcal{S}^*$. This allows us then to define $\vdash^*$, which proves to be an extension of the given $\vdash$.

What is accomplished by such a construction? I assert that it is always possible to expand language and entailment in a coherent fashion so that deductions about the new system will not give us any incorrect conclusions about the old system; moreover a sentential operator $\Rightarrow$ is introduced so that the truth of $(A \Rightarrow B)$ for a valuation of the system is *equivalent* to the holding of $A \vdash^* B$. We could not do this with material implication and it seems to me to be pretty much what Lewis wanted.

Note that, in the construction just outlined, not only do we have the proper rules for $\Rightarrow$ from section IV, but $\Rightarrow$ also satisfies (MP). Consider any such system. Define a unary operation on statements by the equation:

$$\Box A = (T \Rightarrow A)$$

It is quite easy to prove the following:

$$A \Rightarrow B \ \vdash \Box(A \supset B)$$
$$\Box(A \supset B) \ \vdash A \Rightarrow B$$
$$\Box A, \Box(A \supset B) \ \vdash \Box B$$
$$\Box A \ \vdash A$$
$$\Box A \ \vdash \Box\Box A$$

and

$$\frac{\vdash A}{\vdash \Box A}$$

This is nothing more than the Lewis system **S4** with $\Box$ as *necessity* and $\Rightarrow$ as *strict implication*. Actually in the construction of $\mathcal{S}^*$ we will find that $\vdash^*$ satisfies the stronger rules of **S5**; in particular:

$$\vdash \Box A, \Box \sim \Box A$$

A more subtle construction of the extension of the valuations would have avoided this feature, but the details need not concern us here.

I feel that this discussion nearly vindicates Lewis. We have shown how a theory of strict implication can be grafted onto any system of classical propositional calculus. We hasten to add, however, that the extension result does require an *expansion* of the notion of a statement. Quine may wish to object at this point. He might be inclined to say that our $\mathcal{S}^*$ is just a fragment of the metalanguage into which we have *injected* the original object language $\mathcal{S}$. It is just a mathematical trick whereby we allow ourselves to regard $\mathcal{S}^*$ as another object language. (I do not wish to put words into Quine's mouth, but the objection is reasonable.) The answer is, yes, the construction could be viewed this way; but even if one insists on this view, one cannot go on to claim that any act of confusion has been committed. The language $\mathcal{S}^*$ together with its valuations (truth theory) is quite definite, and we can show how to transfer certain results about $\mathcal{S}^*$ back to $\mathcal{S}$. That one may not care to use $\mathcal{S}^*$ for investigations about $\mathcal{S}$ is one's own concern; however, no mistake is involved in making use of $\mathcal{S}^*$.

Whether Lewis would agree to this interpretation, or whether he would feel I have done damage to his program can no longer be determined, and the reader will have to judge for himself. Fortunately we can still ask for Quine's opinion. Lewis himself preferred weaker systems even than S4. The more recent work on semantics for modal logic can be used to understand the formal nature of his systems, but it might be interesting to reexamine his arguments in the light of the result of section IV.

DANA SCOTT

Princeton University