CARNAP AND LOGICAL TRUTH¹)

I

Kant's question 'How are synthetic judgments *a priori* possible?' precipitated the *Critique of Pure Reason*. Question and answer notwithstanding, Mill and others persisted in doubting that such judgments were possible at all. At length some of Kant's own clearest purported instances, drawn from arithmetic, were sweepingly disqualified (or so it seemed; but see § II) by Frege's reduction of arithmetic to logic. Attention was thus forced upon the less tendentious and indeed logically prior question, 'How is logical certainty possible?' It was largely this latter question that precipitated the form of empiricism which we associate with between-war Vienna – a movement which began with Wittgenstein's *Tractatus* and reached its maturity in the work of Carnap.

Mill's position on the second question had been that logic and mathematics were based on empirical generalizations, despite their superficial appearance to the contrary. This doctrine may well have been felt to do less than justice to the palpable surface differences between the deductive sciences of logic and mathematics, on the one hand, and the empirical sciences ordinarily so-called on the other. Worse, the doctrine derogated from the certainty of logic and mathematics; but Mill may not have been one to be excessively disturbed by such a consequence. Perhaps classical mathematics did lie closer to experience then than now; at any rate the infinitistic reaches of set theory, which are so fraught with speculation and so remote from any possible experience, were unexplored in his day. And it is against just these latter-day mathematical extravagances that

¹) This paper was written early in 1954 at the request of Professor Schilpp, for inclusion in a volume on Carnap which he had been planning. The paper has since appeared in Italian translation as 'Carnap e la verità logica', Rivista di Filosofia, vol. 48 (1957), pp. 3–29. Selected portions, running to somewhat less than half, have appeared also in *American Philosophers at Work* (Sidney Hook, ed.), Criterion Books, New York, 1956.

empiricists outside the Vienna Circle have since been known to inveigh,¹) in much the spirit in which the empiricists of Vienna and elsewhere have inveighed against metaphysics.

What now of the empiricist who would grant certainty to logic, and to the whole of mathematics, and yet would make a clean sweep of other nonempirical theories under the name of metaphysics? The Viennese solution of this nice problem was predicated on language. Metaphysics was meaningless through misuse of language; logic was certain through tautologous use of language.

As an answer to the question 'How is logical certainty possible?' this linguistic doctrine of logical truth has its attractions. For there can be no doubt that sheer verbal usage is in general a major determinant of truth. Even so factual a sentence as 'Brutus killed Caesar' owes its truth not only to the killing but equally to our using the component words as we do. Why then should a logically true sentence on the same topic, e.g. 'Brutus killed Caesar or did not kill Caesar', not be said to owe its truth *purely* to the fact that we use our words (in this case 'or' and 'not') as we do? – for it depends not at all for its truth upon the killing.

The suggestion is not, of course, that the logically true sentence is a contingent truth about verbal usage; but rather that it is a sentence which, given the language, automatically becomes true, whereas 'Brutus killed Caeser', given the language, becomes true only contingently on the alleged killing.

Further plausibility accrues to the linguistic doctrine of logical truth when we reflect on the question of alternative logics. Suppose someone puts forward and uses a consistent logic the principles of which are contrary to our own. We are then clearly free to say that he is merely using the familiar particles 'and', 'all', or whatever, in other than the familiar senses, and hence that no real contrariety is present after all. There may of course still be an important failure of intertranslatability, in that the behavior of certain of our logical particles is incapable of being duplicated by paraphrases in his system or vice versa. If the translation in this sense is possible, from his system into ours, then we are pretty sure to protest that he was wantonly using the familiar particles 'and' and

¹) An example is P. W. Bridgman, 'A physicist's second reaction to Mengenlehre,' *Scripta Mathematica*, vol. 2, 1933–4, pp. 101–117, 224–234.

'all' (say) where me might unmisleadingly have used such and such other familiar phrasing. This reflection goes to support the view that the truths of logic have no content over and above the meanings they confer on the logical vocabulary.

Much the same point can be brought out by a caricature of a doctrine of Levy-Bruhl, according to which there are pre-logical peoples who accept certain simple self-contradictions as true. Over-simplifying, no doubt, let us suppose it claimed that these natives accept as true a certain sentence of the form 'p and not p'. Or – not to over-simplify too much – that they accept as true a certain heathen sentence of the form 'q ka bu q' the English translation of which has the form 'p and not p'. But now just how good a translation is this, and what may the lexicographer's method have been? If any evidence can count against a lexicographer's adoption of 'and' and 'not' as translations of 'ka' and 'bu', certainly the natives' acceptance of 'q ka bu q' as true counts overwhelmingly. We are left with the meaninglessness of the doctrine of there being prelogical peoples; prelogicality is a trait injected by bad translators. This is one more illustration of the inseparability of the truths of logic from the meanings of the logical vocabulary.

We thus see that there is something to be said for the naturalness of the linguistic doctrine of logical truth. But before we can get much further we shall have to become more explicit concerning our subject matter.

п

Without thought of any epistemological doctrine, either the linguistic doctrine or another, we may mark out the intended scope of the term 'logical truth', within that of the broader term 'truth', in the following way. First we suppose indicated, by enumeration if not otherwise, what words are to be called logical words; typical ones are 'or', 'not', 'if', 'then', 'and', 'all', 'every', 'only', 'some'. The logical truths, then, are those true sentences which involve only logical words *essentially*. What this means is that any other words, though they may also occur in a logical truth (as witness 'Brutus', 'kill', and 'Caesar' in 'Brutus killed or did not kill Caesar'), can be varied at will without engendering falsity.¹)

¹) Substantially this formulation is traced back a century and a quarter by Yehoshua Bar-Hillel, 'Bolzano's definition of analytic propositions,' *Methodos*, vol. 2, 1950,

Though formulated with reference to language, the above clarification does not of itself hint that logical truths owe their truth to language. What we have thus far is only a delimitation of the class, *per accidens* if you please. Afterward the linguistic doctrine of logical truth, which is an epistemological doctrine, goes on to say that logical truths are true by virtue purely of the intended meanings, or intended usage, of the logical words. Obviously if logical truths *are* true by virtue purely of language, the logical words are the only part of the language that can be concerned in the matter; for these are the only ones that occur essentially.

Elementary logic, as commonly systematized nowadays, comprises truth-function theory, quantification theory, and identity theory. The logical vocabulary for this part, as commonly rendered for technical purposes, consists of truth-function signs (corresponding to 'or', 'and', 'not', etc.), quantifiers and their variables, and '='.

The further part of logic is set theory, which requires there to be classes among the values of its variables of quantification. The one sign needed in set theory, beyond those appropriate to elementary logic, is the connective 'ɛ' of membership. Additional signs, though commonly used for convenience, can be eliminated in well-known ways.

In this dichotomy I leave metatheory, or logical syntax, out of account. For, either it treats of special objects of an extralogical kind, viz. notational expressions, or else, if these are made to give way to numbers by arithmetization, it is reducible via number theory to set theory.

I will not here review the important contrasts between elementary logic and set theory, except for the following one. Every truth of elementary logic is obvious (whatever this really means), or can be made so by some series of individually obvious steps. Set theory, in its present state anyway, is otherwise. I am not alluding here to Gödel's incompleteness principle, but to something right on the surface. Set theory was straining

pp. 32-55 (= *Theoria*, vol. 16, 1950, pp. 91-117). But note that the formulation fails of its purpose unless the phrase 'can be varied at will,' above, is understood to provide for varying the words not only singly but also two or more at at time. E.g., the sentence 'If some men are angels some animals are angels' can be turned into a falsehood by simultaneous substitution for 'men' and 'angels', but not by any substitution for 'angels' alone, nor for 'men', nor for 'animals' (granted the non-existence of angels). For this observation and illustration I am indebted to John R. Myhill, who expresses some indebtedness in turn to Benson Mates. – I added most of this footnote in May, 1955; thus one year after the rest of the essay left my hands.

at the leash of intuition ever since Cantor discovered the higher infinites; and with the added impetus of the paradoxes of set theory the leash was snapped. Comparative set theory has now long been the trend; for, so far as is known, no consistent set theory is both adequate to the purposes envisaged for set theory and capable of substantiation by steps of obvious reasoning from obviously true principles. What we do is develop one or another set theory by obvious reasoning, or elementary logic, from unobvious first principles which are set down, whether for good or for the time being, by something very like convention.

Altogether, the contrasts between elementary logic and set theory are so fundamental that one might well limit the word 'logic' to the former (though I shall not), and speak of set theory as mathematics in a sense exclusive of logic. To adopt this course is merely to deprive ' ϵ ' of the status of a logical word. Frege's derivation of arithmetic would then cease to count as a derivation from logic; for he used set theory. At any rate we should be prepared to find that the linguistic doctrine of logical truth holds for elementary logic and fails for set theory, or vice versa. Kant's readiness to see logic as analytic and arithmetic as synthetic, in particular, is not superseded by Frege's work (as Frege supposed ¹)) if 'logic' be taken as elementary logic. And for Kant logic certainly did not include set theory.

Ш

Where someone disagrees with us as to the truth of a sentence, it often happens that we can convince him by getting the sentence from other sentences, which he does accept, by a series of steps each of which he accepts. Disagreement which cannot be thus resolved I shall call *deductively urresoluble*. Now if we try to warp the linguistic doctrine of logical truth around into something like an experimental thesis, perhaps a first approximation will run thus: *Deductively irresoluble disagreement as to a logical truth is evidence of deviation in usage (or meanings) of words*. This is not yet experimentally phrased, since one term of the affirmed relationship, viz. 'usage' (or 'meanings'), is in dire need of an independent

¹) See §§ 87f., 109 of Gottlob Frege, *Foundations of Arithmetic* (New York: Philosophical Library, and Oxford: Blackwell, 1950), a reprint of *Grundlagen der Arithmetik* (Breslau, 1884) with translation by J. L. Austin.

criterion. However, the formulation would seem to be fair enough within its limits; so let us go ahead with it, not seeking more sublety until need arises.

Already the obviousness or potential obviousness of elementary logic can be seen to present an insuperable obstacle to our assigning any experimental meaning to the linguistic doctrine of elementary logical truth. Deductively irresoluble dissent from an elementary logical truth would count as evidence of deviation over meanings if anything can, but simply because dissent from a logical truism is as extreme as dissent can get.

The philosopher, like the beginner in algebra, works in danger of finding that his solution-in-progress reduces to '0 = 0'. Such is the threat to the linguistic theory of elementary logical truth. For, that theory now seems to imply nothing that is not already implied by the fact that elementary logic is obvious or can be resolved into obvious steps.

The considerations which were adduced in § I, to show the naturalness of the linguistic doctrine, are likewise seen to be empty when scrutinized in the present spirit. One was the circumstance that alternative logics are inseparable practically from mere change in usage of logical words. Another was that illogical cultures are indistinguishable from ill-translated ones. But both of these circumstances are adequately accounted for by mere obviousness of logical principles, without help of a linguistic doctrine of logical truth. For, there can be no stronger evidence of a change in usage than the repudiation of what had been obvious, and no stronger evidence of bad translation than that it translates earnest affirmations into obvious falsehoods.

Another point in § I was that true sentences generally depend for their truth on the traits of their language in addition to the traits of their subject matter; and that logical truths then fit neatly in as the limiting case where the dependence on traits of the subject matter is nil Consider, however, the logical truth 'Everything is self-identical', or '(x) (x = x)'. We can say that it depends for its truth on traits of the language (specifically on the usage of ' = '), and not on traits of its subject matter; but we can also say, alternatively, that it depends on an obvious trait, viz. self-identity, of its subject matter, viz. everything. The tendency of our present reflections is that there is no difference.

I have been using the vaguely psychological word 'obvious' non-technic-

ally, assigning it no explanatory value. My suggestion is merely that the linguistic doctrine of elementary logical truth likewise leaves explanation unbegun. I do not suggest that the linguistic doctrine is false and some doctrine of ultimate and inexplicable insight into the obvious traits of reality is true, but only that there is no real difference between these two pseudo-doctrines.

Turning away now from elementary logic, let us see how the linguistic doctrine of logical truth fares in application to set theory. As noted in § II, we may think of ' ϵ ' as the one sign for set theory in addition to those of elementary logic. Accordingly the version of the linguistic doctrine which was italicized at the beginning of the present section becomes, in application to set theory, this: Among persons already in agreement on elementary logic, deductively irresoluble disagreement as to a truth of set theory is evidence of deviation in usage (or meaning) of ' ϵ '.

This thesis is not trivial in quite the way in which the parallel thesis for elementary logic was seen to be. It is not indeed experimentally significant as it stands, simply because of the lack, noted earlier, of a separate criterion for usage or meaning. But it does seem reasonable, by the following reasoning.

Any acceptable evidence of usage or meaning of words must reside surely either in the observable circumstances under which the words are uttered (in the case of concrete terms referring to observable individuals) or in the affirmation and denial of sentences in which the words occur. Only the second alternative is relevant to 'ɛ'. Therefore any evidence of deviation in usage or meaning of 'ɛ' must reside in disagreement on sentences containing 'ɛ'. This is not, of course, to say of every sentence containing 'ɛ' that disagreement over it establishes deviation in usage or meaning of 's'. We have to assume in the first place that the speaker under investigation agrees with us on the meanings of words other than ε' in the sentences in question. And it might well be that, even from among the sentences containing only 'ɛ' and words on whose meanings he agrees with us, there is only a select species S which is so fundamental that he cannot dissent from them without betraying deviation in his usage or meaning of 'ɛ'. But S may be expected surely to include some (if not all) of the sentences which contain nothing but 's' and the elementary logical particles; for it is these sentences, insofar as true, that constitute (pure, or unapplied) set theory. But it is difficult to conceive of how to be

other than democratic toward the truths of set theory. In exposition we may select some of these truths as so-called postulates and deduce others from them, but this is subjective discrimination, variable at will, expository and not set-theoretic. We do not change our meaning of ' ε ' between the page where we show that one particular truth is deducible by elementary logic from another and the page where we show the converse. Given this democratic outlook, finally, the law of sufficient reason leads us to look upon S as including *all* the sentences which contain only ' ε ' and the elementary logic and in irresoluble disagreement on set theory is in deviation with respect to the usage or meaning of ' ε '; and this was the thesis.

The effect of our effort to inject content into the linguistic doctrine of logical truth has been, up to now, to suggest that the doctrine says nothing worth saying about elementary logical truth, but that when applied to set-theoretic truth it makes for a reasonable partial condensation of the otherwise vaporous notion of meaning as applied to ' ε '.

IV

The linguistic doctrine of logical truth is sometimes expressed by saying that such truths are true by linguistic convention. Now if this be so, certainly the conventions are not in general explicit. Relatively few persons, before the time of Carnap, had ever seen any convention that engendered truths of elementary logic. Nor can this circumstance be ascribed merely to the slipshod ways of our predecessors. For it is impossible in principle, even in an ideal state, to get even the most elementary part of logic exclusively by the explicit application of conventions stated in advance. The difficulty is the vicious regress, familiar from Lewis Carroll,¹) which I have elaborated elsewhere.²) Briefly the point is that the logical truths, being infinite in number, must be given by general conventions rather than singly; and logic is needed then to begin with, in the metatheory, in order to apply the general conventions to individual cases.

¹⁾ What the tortoise said to Achilles,' Mind, vol. 4, 1895, pp. 278ff.

²) 'Truth by convention,' in O. H. Lee (ed.), *Philosophical Essays for A. N. Whitehead* (New York, 1936), pp. 90–124. Reprinted in H. Feigl and W. Sellars (eds.), *Readings in Philosophical Analysis* (New York: Appleton, 1949).

'In dropping the attributes of deliberateness and explicitness from the notion of linguistic convention,' I went on to complain in the aforementioned paper, 'we risk depriving the latter of any explanatory force and reducing it to an idle label.' It would seem that to call elementary logic true by convention is to add nothing but a metaphor to the linguistic doctrine of logical truth which, as applied to elementary logic, has itself come to seem rather an empty figure (cf. § III).

The case of set theory, however, is different on both counts. For set theory the linguistic doctrine has seemed less empty (cf. § III); in set theory, moreover, convention in quite the ordinary sense seems to be pretty much what goes on (cf. § II). Conventionalism has a serious claim to attention in the philosophy of mathematics, if only because of set theory. Historically, though, conventionalism was encouraged in the philosophy of mathematics rather by the non-Euclidean geometries and abstract algebras, with little good reason. We can contribute to subsequent purposes by surveying this situation. Further talk of set theory is deferred to § V.

In the beginning there was Euclidean geometry, a compendium of truths about form and void; and its truths were not based on convention (except as a conventionalist might, begging the present question, apply this tag to everything mathematical). Its truths were in practice presented by deduction from so-called postulates (including axioms; I shall not distinguish); and the selection of truths for this role of postulate, out of the totality of truths of Euclidean geometry, was indeed a matter of convention. But this is not *truth* by convention. The truths were there, and what was conventional was merely the separation of them into those to be taken as starting point (for purposes of the exposition at hand) and those to be deduced from them.

The non-Euclidean geometries came of artificial deviations from Euclid's postulates, without thought (to begin with) of true interpretation. These departures were doubly conventional; for Euclid's postulates were a conventional selection from among the truths of geometry, and then the departures were arbitrarily or conventionally devised in turn. But still there was no truth by convention, because there was no truth.

Playing within a non-Euclidean geometry, one might conveniently make believe that his theorems were interpreted and true; but even such conventional make-believe is not truth by convention. For it is not really truth at all; and what is conventionally pretended is that the theorems are true by non-convention.

Non-Euclidean geometries have, in the fullness of time, received serious interpretations. This means that ways have been found of so construing the hitherto unconstrued terms as to identify the at first conventionally chosen set of non-sentences with some genuine truths, and truths presumably not by convention. The status of an interpreted non-Euclidean geometry differs in no basic way from the original status of Euclidean geometry, noted above.

Uninterpreted systems became quite the fashion after the advent of non-Euclidean geometries. This fashion helped to cause, and was in turn encouraged by, an increasingly formal approach to mathematics. Methods had to become more formal to make up for the unavailability, in uninterpreted systems, of intuition. Conversely, disinterpretation served as a crude but useful device (until Frege's syntactical approach came to be appreciated) for achieving formal rigor uncorrupted by intuition.

The tendency to look upon non-Euclidean geometries as true by convention applied to uninterpreted systems generally, and then carried over from these to mathematical systems generally. A tendency indeed developed to look upon all mathematical systems as, qua mathematical, uninterpreted. This tendency can be accounted for by the increase of formality, together with the use of disinterpretation as a heuristic aid to formalization. Finally, in an effort to make some sense of mathematics thus drained of all interpretation, recourse was had to the shocking quibble of identifying mathematics merely with the elementary logic which leads from uninterpreted postulates to uninterpreted theorems.¹) What is shocking about this is that it puts arithmetic qua interpreted theory of number, and analysis qua interpreted theory of functions, and geometry qua interpreted theory of space, outside mathematics altogether.

The substantive reduction of mathematics to logic by Frege, Whitehead, and Russell is of course quite another thing. It is a reduction not to elementary logic but to set theory; and it is a reduction of genuine interpreted mathematics, from arithmetic onward.

¹) Bertrand Russell, *Principles of Mathematics* (Cambridge, 1903), pp. 429f.; Heinrich Behmann, 'Sind die mathematischen Urteile Analytisch oder synthetisch?' *Erkenntnis*, vol. 4, 1934, pp. 8ff.; and others.

V

Let us then put aside these confusions and get back to set theory. Set theory is pursued as interpreted mathematics, like arithmetic and analysis; indeed, it is to set theory that those further branches are reducible. In set theory we discourse about certain immaterial entities, real or erroneously alleged, viz. sets, or classes. And it is in the effort to make up our minds about genuine truth and falsity of sentences about these objects that we find ourselves engaged in something very like convention in an ordinary non-metaphorical sense of the word. We find ourselves making deliberate choices and setting them forth unaccompanied by any attempt at justification other than in terms of elegance and convenience. These adoptions, called postulates, and their logical consequences (via elementary logic), are true until further notice.

So here is a case where postulation can plausibly be looked on as constituting truth by convention. But in § IV we have seen how the philosophy of mathematics can be corrupted by supposing that postulates always play that role. Insofar as we would epistemologize and not just mathematize, we might divide postulation as follows. Uninterpreted postulates may be put aside, as no longer concerning us; and on the interpreted side we may distinguish between legislative and discursive postulation. Legislative postulation institutes truth by convention, and seems plausibly illustrated in contemporary set theory. On the other hand discursive postulation is mere selection, from a preëxisting body of truths, of certain ones for use as a basis from which to derive others, initially known or unknown. What discursive postulation fixes is not truth, but only some particular ordering of the truths, for purposes perhaps of pedagogy or perhaps of inquiry into logical relationships ('logical' in the sense of elementary logic). All postulation is of course conventional, but only legislative postulation properly hints of *truth* by convention.

It is well to recognize, if only for its distinctness, yet a further way in which convention can enter; viz., in the adoption of new notations for old ones, without, as one tends to say, change of theory. Truths containing the new notation are conventional transcriptions of sentences true apart from the convention in question. They depend for their truth partly on language, but then so did 'Brutus killed Caesar' (cf. § I). They come into being through a conventional adoption of a new sign, and they

become true through conventional definition of that sign *together with* whatever made the corresponding sentences in the old notation true.

Definition, in a properly narrow sense of the word, is convention in a properly narrow sense of the word. But the phrase 'true by definition' must be taken cautiously; in its strictest usage it refers to a transcription, by the definition, of a truth of elementary logic. Whether such a sentence is true by convention depends on whether the logical truths themselves be reckoned as true by convention. Even an outright equation or biconditional connection the definiens and the definiendum is a definitional transcription of a prior logical truth of the form 'x = x' or ' $p \equiv p$ '.

Definition commonly so-called is not thus narrowly conceived, and must for present purposes be divided, as postulation was divided, into legislative and discursive. Legislative definition introduces a notation hitherto unused, or used only at variance with the practice proposed, or used also at variance, so that a convention is wanted to settle the ambiguity. Discursive definition, on the other hand, sets forth a preëxisting relation of interchangeability or coextensiveness between notations in already familiar usage. A frequent purpose of this activity is to show how some chosen part of language can be made to serve the purposes of a wider part. Another frequent purpose is language instruction.

It is only legislative definition, and not discursive definition nor discursive postulation, that makes a conventional contribution to the truth of sentences. Legislative postulation, finally, affords truth by convention unalloyed.

Increasingly the word 'definition' connotes the formulas of definition which appear in connection with formal systems, signalled by some extrasystematic sign such as ' $=_{df}$ '. Such definitions are best looked upon as correlating two systems, two notations, one of which is prized for its economical lexicon and the other for its brevity or familiarity of expression.¹) Definitions so used can be either legislative or discursive in their inception. But this distinction is in practice left unindicated, and wisely; for it is a distinction only between particular acts of definition, and not germane to the definition as an enduring channel of intertranslation.

The distinction between the legislative and the discursive refers thus to

1) See my From a Logical Point of View (Cambridge, Mass.: Harvard, 1953), pp. 26f.

the act, and not to its enduring consequence, in the case of postulation as in the case of definition. This is because we are taking the notion of truth by convention fairly literally and simple-mindedly, for lack of an intelligible alternative. So conceived, conventionality is a passing trait, significant at the moving front of science but useless in classifying the sentences behind the lines. It is a trait of events and not of sentences.

Might we not still project a derivative trait upon the sentences themselves, thus speaking of a sentence as forever true by convention if its first adoption as true was a convention? No; this, if done seriously, involves us in the most unrewarding historical conjecture. Legislative postulation contributes truths which become integral to the corpus of truths; the artificiality of their origin does not linger as a localized quality, but suffuses the corpus. If a subsequent expositor singles out those once legislatively postulated truths again as postulates, this signifies nothing; he is engaged only in discursive postulation. He could as well choose his postulates from elsewhere in the corpus, and will if he thinks this serves his expository ends.

VI

Set theory, currently so caught up in legislative postulation, may some day gain a norm – even a strain of obviousness, perhaps – and lose all trace of the conventions in its history. A day could likewise have been when our elementary logic was itself instituted as a deliberately conventional deviation from something earlier, instead of evolving, as it did, mainly by unplanned shifts of form and emphasis coupled with casual novelties of notation.

Today indeed there are dissident logicians even at the elementary level, propounding deviations from the law of the excluded middle. These deviations, insofar as meant for serious use and not just as uninterpreted systems, are as clear cases of legislative postulation as the ones in set theory. For here we have again, quite as in set theory, the propounding of a deliberate choice unanccompanied (conceivably) by any attempt at justification other than in terms of convenience.

This example from elementary logic controverts no conclusion we have reached. According to §§ II and III, the departure from the law of the excluded middle would count as evidence of revised usage of 'or' and 'not'. (This judgment was upheld in § III, though disqualified as evidence

for the linguistic doctrine of logical truth.) For the deviating logician the words 'or' and 'not' are unfamiliar, or defamiliarized; and his decisions regarding truth values for their proposed contexts can then be just as genuinely a matter of deliberate convention as the decisions of the creative set theorist regarding contexts of ' ϵ '.

The two cases are indeed much alike. Not only is departure from the classical logic of 'or' and 'not' evidence of revised usage of 'or' and 'not'; likewise, as argued at length in § III, divergences between set-theorists may reasonably be reckoned to revised usage of ' ϵ '. Any such revised usage is conspicuously a matter of convention, and can be declared by legislative postulation.

We have been at a loss to give substance to the linguistic doctrine, particularly of elementary logical truth, or to the doctrine that the familiar truths of logic are true by convention. We have found some sense in the notion of truth by convention, but only as attaching to a process of adoption, viz. legislative postulation, and not as a significant lingering trait of the legislatively postulated sentence. Surveying current events, we note legislative postulation in set theory and, at a more elementary level, in connection with the law of the excluded middle.

And do we not find the same continually in the theoretical hypotheses of natural science itself? What seemed to smack of convention in set theory (§ V), at any rate, was 'deliberate choice, set forth unaccompanied by any attempt at justification other than in terms of elegance and convenience'; and to what theoretical hypothesis of natural science might not this same character be attributed? For surely the justification of any theoretical hypothesis can, at the time of hypothesis, consist in no more than the elegance or convenience which the hypothesis brings to the containing body of laws and data. How then are we to delimit the category of legislative postulation, short of including under it every new act of scientific hypothesis?

The situation may seem to be saved, for ordinary hypotheses in natural science, by there being some indirect but eventual confrontation with empirical data. However, this confrontation can be remote; and, conversely, some such remote confrontation with experience may be claimed even for pure mathematics and elementary logic. The semblance of a difference in this respect is largely due to over-emphasis of departmental boundaries. For a self-contained theory which we can check with

experience includes, in point of fact, not only its various theoretical hypotheses of so-called natural science but also such portions of logic and mathematics as it makes use of. Hence I do not see how a line is to be drawn between hypotheses which confer truth by convention and hypotheses which do not, short of reckoning *all* hypotheses to the former category save perhaps those actually derivable or refutable by elementary logic from what Carnap used to call protocol sentences. But this version, besides depending to an unwelcome degree on the debatable notion of protocol sentences, is far too inclusive to suit anyone.

Evidently our troubles are waxing. We had been trying to make sense of the role of convention in *a priori* knowledge. Now the very distinction between *a priori* and empirical begins to waver and dissolve, at least as a distinction between sentences. (It could of course still hold as a distinction between factors in one's adoption of a sentence, but both factors might be operative everywhere.)

VII

Whatever our difficulties over the relevant distinctions, it must be conceded that logic and mathematics do seem qualitatively different from the rest of science. Logic and mathematics hold conspicuously aloof from any express appeal, certainly, to observation and experiment. Having thus nothing external to look to, logicians and mathematicians look closely to notation and explicit notational operations: to expressions, terms, substitution, transposition, cancellation, clearing of fractions, and the like. This concern of logicians and mathematicians with syntax (as Carnap calls it) is perennial, but in modern times it has become increasingly searching and explicit, and has even prompted, as we see, a linguistic philosophy of logical and mathematical truth.

On the other hand an effect of these same formal developments in modern logic, curiously, has been to show how to divorce mathematics (other than elementary logic) from any peculiarly notational considerations not equally relevant to natural science. By this I mean that mathematics can be handled (insofar as it can be handled at all) by axiomatization, outwardly quite like any system of hypotheses elsewhere in science; and elementary logic can then be left to extract the theorems.

The consequent affinity between mathematics and systematized natural

science was recognized by Carnap when he propounded his P-rules alongside his L-rules or meaning postulates. Yet he did not look upon the P-rules as engendering analytic sentences, sentences true purely by language. How to sustain this distinction has been very much our problem in these pages, and one on which we have found little encouragement.

Carnap appreciated this problem, in *Logical Syntax*, as a problem of finding a difference in kind between the P-rules (or the truths thereby specified) and the L-rules (or the L-truths, analytic sentences, thereby specified). Moreover he proposed an ingenious solution.¹) In effect he characterized the logical (including mathematical) vocabulary as the largest vocabulary such that (1) there are sentences which contain only that vocabulary and (2) all such sentences are determinable as true or false by a purely syntactical condition – i.e., by a condition which speaks only of concatenation of marks. Then he limited the L-truths in effect to those involving just the logical vocabulary essentially.²)

Truths given by P-rules were supposedly excluded from the category of logical truth under this criterion, because, though the rules specifying them are formally stated, the vocabulary involved can also be recombined to give sentences whose truth values are not determinate under any set of rules formally formulable in advance.

At this point one can object (pending a further expedient of Carnap's, which I shall next explain) that the criterion based on (1) and (2) fails of its purpose. For, consider to begin with the totality of those sentences which are expressed purely within what Carnap (or anyone) would want to count as logical (and mathematical) vocabulary. Suppose, in conformity with (2), that the division of this totality into the true and the false is reproducible in purely syntactical terms. Now surely the adding of one general term of an extra-logical kind, say 'heavier than', is not going to alter the situation. The truths which are expressible in terms of just 'heavier than', together with the logical vocabulary, will be truths of only the most general kind, such as $(\exists x) (\exists y) (x \text{ is heavier than } y)$ ', $(x) \sim (x \text{ is}$ heavier than x)', and $(x)(y)(z)(x \text{ is heavier than } y \cdot y \text{ is heavier than } z \cdot \Im$. x is heavier than z)'. The division of the truths from the falsehoods

¹) Carnap, Logical Syntax of Language, § 50.

²) Cf. § I above. Also, for certain reservations conveniently postponed at the moment, see § IX on 'essential predication.'

in this supplementary domain can probably be reproduced in syntactical terms if the division of the original totality could. But then, under the criterion based on (1) and (2), 'heavier than' qualifies for the logical vocabulary. And it is hard to see what whole collection of general terms of natural science might not qualify likewise.

The further expedient, by which Carnap met this difficulty, was his use of Cartesian coördinates.¹) Under this procedure, each spatio-temporal particular c becomes associated with a class K of quadruples of real numbers, viz., the class of those quadruples which are the coördinates of component point-events of c. Further let us write K[t] for the class of triples which with t appended belong to K; thus K[t] is that class of triples of real numbers which is associated with the momentary state of object c at time t. Then, in order to say e.g. that c_1 is heavier than c_2 at time t, we say ' $H(K_1[t], K_2[t])$ ', which might be translated as 'The momentary object associated with $K_1[t]$ is heavier than that associated with $K_2[t]$.' Now $K_1[t]$ and $K_2[t]$ are, in every particular case, purely mathematical objects; viz., classes of triples of real numbers. So let us consider all the true and false sentences of the form $H(K_1[t], K_2[t])$ where, in place of $K_1[t]$ and $K_2[t]$, we have purely logico-mathematical designations of particular classes of triples of real numbers. There is no reason to suppose that all the truths of this domain can be exactly segregated in purely syntactical terms. Thus inclusion of 'H' does violate (2), and therefore 'H' fails to qualify as logical vocabulary. By adhering to the method of coördinates and thus reconstruing all predicates of natural science in the manner here illustrated by 'H', Carnap overcomes the objection noted in the preceding paragraph.

To sum up very roughly, this theory characterizes logic (and mathematics) as the largest part of science within which the true-false dichotomy *can* be reproduced in syntactical terms. This version may seem rather thinner than the claim that logic and mathematics are somehow true by linguistic convention, but at any rate it is more intelligible, and, if true, perhaps interesting and important. To become sure of its truth, interest, and importance, however, we must look more closely at this term 'syntax'.

As used in the passage: 'The terms 'sentence' and 'direct consequence' are the two primitive terms of logical syntax,'2) the term 'syntax' is of

²) Carnap, Philosophy and Logical Syntax, p. 47.

¹⁾ Logical Syntax of Language, §§ 3, 15.

course irrelevant to a thesis. The relevant sense is that rather in which it connotes discourse about marks and their succession. But here still we must distinguish degrees of inclusiveness; two different degrees are exemplified in Logical Syntax, according as the object language is Carnap's highly restricted Language I or his more powerful Language II. For the former, Carnap's formulation of logical truth is narrowly syntactical in the manner of familiar formalizations of logical systems by axioms and rules of inference. But Gödel's proof of the incompletability of elementary number theory shows that no such approach can be adequate to mathematics in general, nor in particular to set theory, nor to Language II. For Language II, in consequence, Carnap's formulation of logical truth proceeded along the lines rather of Tarski's technique of truth-definition.¹) The result was still a purely syntactical specification of the logical truths, but only in this more liberal sense of 'syntactical': it was couched in a vocabulary consisting (in effect) of (a) names of signs, (b) an operator expressing concatenation of expressions, and (c), by way of auxiliary machinery, the whole logical (and mathematical) vocabulary itself.

So construed, however, the thesis that logico-mathematical truth is syntactically specifiable becomes uninteresting. For, what it says is that logico-mathematical truth is specifiable in a notation consisting solely of (a), (b), and the whole logico-mathematical vocabulary itself. But *this* thesis would hold equally if 'logico-mathematical' were broadened (at *both* places in the thesis) to include physics, economics, and anything else under the sun; Tarski's routine of truth-definition would still carry through just as well. No special trait of logic and mathematics has been singled out after all.

Strictly speaking, the position is weaker still. The mathematics appealed to in (c) must, as Tarski shows, be a yet more inclusive mathematical theory in certain respects than that for which truth is being defined. It was largely because of his increasing concern over this self-stultifying situation that Carnap relaxed his stress on syntax, in the years following *Logical Syntax*, in favor of semantics.

¹) Logical Syntax, especially §§ 34a-i, 60a-d, 71a-d. These sections had been omitted from the German edition, but only for lack of space; cf. p. xi of the English edition. Meanwhile they had appeared as articles: 'Die Antinomien...' and 'Ein Gültigkeits-kriterium...' At that time Carnap had had only partial access to Tarski's ideas

VIII

Even if logical truth were specifiable in syntactical terms, this would not show that it was grounded in language. Any *finite* class of truths (to take an extreme example) is clearly reproducible by a membership condition couched in as narrowly syntactical terms as you please; yet we certainly cannot say of every finite class of truths that its members are true purely by language. Thus the ill-starred doctrine of syntactical specifiability of logical truth was always something other than the linguistic doctrine of logical truth, if this be conceived as the doctrine that logical truth is grounded in language. In any event the doctrine of syntactical specifiability, which we found pleasure in being able to make comparatively clear sense of, has unhappily had to go by the board. The linguistic doctrine of logical truth, on the other hand, goes sturdily on.

The notion of logical truth is now counted by Carnap as semantical. This of course does not of itself mean that logical truth is grounded in language; for note that the general notion of truth is also semantical, though truth in general is not grounded purely in language. But the semantical attribute of logical truth, in particular, *is* one which, according to Carnap, is grounded in language: in convention, fiat, meaning. Such support as he hints for this doctrine, aside from ground covered in §§ I-VI, seems to depend on an analogy with what goes on in the propounding of artificial languages; and I shall now try to show why I think the analogy mistaken.

I may best schematize the point by considering a case, not directly concerned with logical truth, where one might typically produce an artificial language as a step in an argument. This is the imaginary case of a logical positivist, say Ixmann, who is out to defend scientists against the demands of a metaphysician. The metaphysician argues that science presupposes metaphysical principles, or raises metaphysical problems, and that the scientists should therefore show due concern. Ixmann's answer consists in showing in detail how people (on Mars, say) might speak a language quite adequate to all of our science but, unlike our language, incapable of expressing the alleged metaphysical issues. (I applaud

⁽cf. 'Gültigkeitskriterium,' footnote 3), the full details of which reached the non-Slavic world in 1936: Alfred Tarski, 'Der Wahrheitsbegriff in den formalisierten Sprachen,' *Studia Philosophica*, vol. 1, pp. 261-405.

this answer, and think it embodies the most telling component of Carnap's own anti-metaphysical representations; but here I digress.) Now how does our hypothetical Ixmann specify that doubly hypothetical language? By telling us, at least to the extent needed for his argument, what these Martians are to be imagined as uttering and what they are thereby to be understood to mean. Here is Carnap's familiar duality of formation rules and transformation rules (or meaning postulates), as rules of language. But these rules are part only of Ixmann's narrative machinery, not part of what he is portraying. He is not representing his hypothetical Martians themselves as somehow explicit on formation and transformation rules. Nor is he representing there to be any intrinsic difference between those truths which happen to be disclosed to us by his partial specifications (his transformation rules) and those further truths, hypothetically likewise known to the Martians of his parable, which he did not trouble to sketch in.

The threat of fallacy lurks in the fact that Ixmann's rules are indeed arbitrary flats, as is his whole Martian parable. The fallacy consists in confusing levels, projecting the conventional character of the rules into the story, and so misconstruing Ixmann's parable as attributing truthlegislation to his hypothetical Martians.

The case of a non-hypothetical artificial language is in principle the same. Being a new invention, the language has to be explained; and the explanation will proceed by what may certainly be called formation and transformation rules. These rules will hold by arbitrary fiat, the artifex being boss. But all we can reasonably ask of these rules is that they enable us to find corresponding to each of his sentences a sentence of like truth value in familiar ordinary language. There is no (to me) intelligible additional decree that we can demand of him as to the boundary between analytic and synthetic, logic and fact, among his truths. We may well decide to extend our word 'analytic' or 'logically true' to sentences of his language which he in his explanations has paired off fairly directly with English sentences so classified by us; but this is our decree, regarding our word 'analytic' or 'logically true'.

We had in § II to form some rough idea of what logical truth was supposed

to take in, before we could get on with the linguistic doctrine of logical truth. This we did, with help of the general notion of truth ¹) together with a partial enumeration of the logical vocabulary of a particular language. In § VII we found hope of a less provincial and accidental characterization of logical vocabulary; but it failed. Still, the position is not intolerable. We well know from modern logic how to devise *a* technical notation which is admirably suited to the business of 'or', 'not', 'and', 'all,' 'only', and such other particles as we would care to count as logical; and to enumerate the signs and constructions of that technical notation, or a theoretically adequate subset of them, is the work of a moment (cf. § II). Insofar as we are content to think of all science as fitted within that stereotyped logical framework – and there is no hardship in so doing – our notion of logical truth. But only in point of extent. There is no epistemological corollary as to the *ground* of logical truth (cf. § II).

Even this half-way tolerable situation obtains only for logical truth in a relatively narrow sense, omitting truths by 'essential predication' (in Mill's phrase) such as 'No bachelor is married'.²) I tend to reserve the term 'logically true' for the narrower domain, and to use the term 'analytic' for the more inclusive domain which includes truths by essential predication. Carnap on the contrary has used both terms in the broader sense. But the problems of the two subdivisions of the analytic class differ in such a way that it has been convenient up to now in this essay to treat mainly of logical truth in the narrower sense.

The truths by essential predication are sentences which can be turned into logical truths by supplanting certain simple predicates (e.g. 'bachelor') by complex synonyms (e.g. 'man not married'). This formulation is not inadequate to such further examples as 'If A is part of B and B is part of C then A is part of C'; this case can be managed by using for 'is part of' the synonym 'overlaps nothing save what overlaps'.³) The relevant notion

¹) In defense of this general notion, in invidious contrast to that of analyticity, see my *From a Logical Point of View*, pp. 137f.

²⁾ Cf. M. White, 'The analytic and the synthetic: an untenable dualism,' in Sidney Hook (ed.), *John Dewey: Philosopher of Science and Freedom* (New York: Dial, 1950), pp. 316–330. Reprinted in Leonard Linsky (ed.), *Semantics and the Philosophy of Language* (Urbana: University of Illinois Press, 1952).

³) Cf. Nelson Goodman, *The Structure of Appearance* (Cambridge, Mass.: Harvard, 1951).

of synonymy is simply *analytic* coextensiveness (however circular this might be as a definition).

To count analyticity a genus of logical truth is to grant, it may seem, the linguistic doctrine of logical truth; for the term 'analytic' directly suggests truth by language. But this suggestion can be adjusted, in parallel to what was said of 'true by definition' in \S V. 'Analytic' means true by synonymy and logic, hence no doubt true by language and logic, and simply true by language *if* the linguistic doctrine of logical truth is right. Logic itself, throughout these remarks, may be taken as including or excluding set theory (and hence mathematics), depending on further details of one's position.

What has made it so difficult for us to make satisfactory sense of the linguistic doctrine is the obscurity of 'true by language'. Now 'synonymous' lies within that same central obscurity; for, about the best we can say of synonymous predicates is that they are somehow 'coextensive by language'. The obscurity extends, of course, to 'analytic'.

One quickly identifies certain seemingly transparent cases of synonymy, such as 'bachelor' and 'man not married', and senses the triviality of associated sentences such as 'No bachelor is married'. Conceivably the mechanism of such recognition, when better understood, might be made the basis of a definition of synonymy and analyticity in terms of linguistic behavior. On the other hand such an approach might make sense only of something like degrees of synonymy and analyticity. I see no reason to expect that the full-width analyticity which Carnap and others make such heavy demands upon can be fitted to such a foundation in even an approximate way. In any event, we at present lack any tenable general suggestion, either rough and practical or remotely theoretical, as to what it is to be an analytic sentence. All we have are purported illustrations, and claims that the truths of elementary logic, with or without the rest of mathematics, should be counted in. Wherever there has been a semblance of a general criterion, to my knowledge, there has been either some drastic failure such as tended to admit all or no sentences as analytic, or there has been a circularity of the kind noted three paragraphs back, or there has been a dependence on terms like 'meaning', 'possible', 'conceivable', and the like, which are at least as mysterious (and in the same way) as what we want to define. I have expatiated on these troubles elsewhere, as has White.1)

1) Quine, From a Logical Point of View, Essay II; White, op. cit.

Logical truth (in my sense, excluding the additional category of essential predication) is, we saw, well enough definable (relatively to a fixed logical notation). *Elementary* logical truth can even be given a narrowly syntactical formulation, such as Carnap once envisaged for logic and mathematics as a whole (cf. § VII); for the deductive system of elementary logic is known to be complete. But when we would supplement the logical truths by the rest of the so-called analytic truths, true by essential predication, then we are no longer able even to say what we are talking about. The distinction itself, and not merely an epistemological question concerning it, is what is then in question.

What of settling the limits of the broad class of analytic truths by fixing on a standard language as we did for logical truth? No, the matter is very different. Once given the logical vocabulary, we have a means of clearly marking off the species logical truth within the genus truth. But the intermediate genus analyticity is not parallel, for it does not consist of the truths which contain just a certain vocabulary essentially (in the sense of § II). To segregate analyticity we should need rather some sort of accounting of synonymies throughout a universal language. No regimented universal language is at hand, however, for adoption or consideration; what Carnap has propounded in this direction have of course been only illustrative samples, fragmentary in scope. And even if there were one, it is not clear by what standards we would care to settle questions of synonymy and analyticity within it.

Х

Carnap's present position 1) is that one has specified a language quite rigorously only when he has fixed, by dint of so-called meaning postulates, what sentences are to count as analytic. The proponent is supposed to distinguish between those of his declarations which count as meaning postulates, and thus engender analyticity, and those which do not. This he does, presumably, by attaching the label 'meaning postulate'.

But the sense of this label is far less clear to me than four causes of its seeming to be clear. Which of these causes has worked on Carnap, if any, I cannot say; but I have no doubt that all four have worked on his readers. One of these causes is misevaluation of the role of convention in con-¹) See particularly 'Meaning postulates.'

nection with artificial language; thus note the unattributed fallacy described in § VIII. Another is misevaluation of the conventionality of postulates: failure to appreciate that postulates, though they are postulates always by fiat, are not *therefore* true by fiat; cf. §§ IV-V. A third is over-estimation of the distinctive nature of postulates, and of definitions, because of conspicuous and peculiar roles which postulates and definitions have played in situations not really relevant to present concerns: postulates in uninterpreted systems (cf. § IV), and definitions in double systems of notation (cf. § V). A fourth is misevaluation of legislative postulation and legislative definition themselves, in two respects: failure to appreciate that this legislative trait is a trait of scientific hypotheses very generally (cf. § VI), and failure to appreciate that it is a trait of the passing event rather than of the truth which is thereby instituted (cf. end of § V).

Suppose a scientist introduces a new term, for a certain substance or force. He introduces it by an act either of legislative definition or of legislative postulation. Progressing he evolves hypotheses regarding further traits of the named substance or force. Suppose now that some such eventual hypothesis, well attested, identifies this substance or force with one named by a complex term built up of other portions of his scientific vocabulary. We all know that this new identity will figure in the ensuing developments quite on a par with the identity which first came of the act of legislative definition, if any, or on a par with the law which first came of the act of legislative postulation. Revisions, in the course of further progress, can touch any of these affirmations equally. Now I urge that scientists, proceeding thus, are not thereby slurring over any meaningful distinction. Legislative acts occur again and again; on the other hand a dichotomy of the resulting truths themselves into analytic and synthetic, truths by meaning postulate and truths by force of nature, has been given no tolerably clear meaning even as a methodological ideal. One conspicuous consequence of Carnap's belief in this dichotomy may be seen in his attitude toward philosophical issues, e.g. as to what there is. It is only by assuming the cleavage between analytic and synthetic truths that he is able e.g. to declare the problem of universals to be a matter not of theory but of linguistic decision.¹) Now I am as impressed as anyone

¹) See Carnap, 'Empiricism, semantics, and ontology,' Revue internationale de Philosophie, 1950, especially § 3, longest footnote.

with the vastness of what language contributes to science and to one's whole view of the world; and in particular I grant that one's hypothesis as to what there is, e.g. as to there being universals, is at bottom just as arbitrary or pragmatic a matter as one's adoption of a new brand of set theory or even a new system of bookkeeping. Carnap in turn recognizes that such decisions, however conventional, 'will nevertheless usually be influenced by theoretical knowledge.'¹) But what impresses me more than it does Carnap is how well this whole attitude is suited also to the theoretical hypotheses of natural science itself, and how little basis there is for a distinction.

The lore of our fathers is a fabric of sentences. In our hands it develops and changes, through more or less arbitrary and deliberate revisions and additions of our own, more or less directly occasioned by the continuing stimulation of our sense organs. It is a pale grey lore, black with fact and white with convention. But I have found no substantial reasons for concluding that there are any quite black threads in it, or any white ones.

Harvard University

1) Op cit., § 2, fifth paragraph.