VOLUME LXII, No. 10 MAY 13, 1965

THE JOURNAL OF PHILOSOPHY

CRAIG'S THEOREM

Craig's Observation. "Craig's theorem," * as philosophers call it, is actually a corollary to an observation. The observation is that (I) *Every theory that admits a recursively enumerable set of axioms can be recursively axiomatized.*

Some explanations are in order here: (1) A *theory* is an infinite set of wffs (well-formed formulas) which is closed under the usual rules of deduction. One way of giving a theory T is to specify a set of sentences S (called the *axioms of* T) and to define T to consist of the sentences S together with all sentences that can be derived from (one or more) sentences in S by means of logic. (2) If T is a theory with axioms S, and S' is a subset of T such that every member of S can be deduced from sentences in S', then S' is called an *alternative set of axioms for* T. Every theory admits of infinitely many alternative axiomatizations—including the trivial axiomatization, in which every member of T is taken as an axiom (i.e., $S = T$). (3) A set S is called *recursive* if and only if it is decidable—i.e., there exists an effective procedure for telling whether or not an arbitrary wff belongs to S. (This is not the mathematical definition of 'recursive', of course, but the intuitive concept which the mathematical definition captures.) For 'effective procedure' one can also write 'Turing machine'.[1]

A theory is *recursively axiomatizable* (often simply "axiomatizable," in the literature) if it has at least one set of axioms that is recursive. Every finite set is recursive; thus all theories that can be finitely axiomatized are recursively axiomatizable. An example of a theory that can be recursively axiomatized but not finitely axiomatized is Peano arithmetic. The primitive predicates are $\mathrm{E}(x,y)$ (also written $x = y$), $\mathrm{S}(x,y,z)$ (also written $x + y = z$), $\mathrm{T}(x,y,z)$ (also written $xy = z$), and $\mathrm{F}(x,y)$ (also written $y = x'$ or $y = x + 1$). The axioms are Peano's axioms for number theory

* William Craig, "On Axiomatizability within a System," *Journal of Symbolic Logic*, **18**, 1 (March, 1953): 30–32. See also "Replacement of Auxiliary Expressions," *Philosophical Review*, **55**, 1 (January, 1956): 38–55.

[1] Cf. Martin Davis, *Computability and Unsolvability* (New York: McGraw-Hill, 1958).

plus the four formulas that recursively define addition and multiplication $[(x)(x+0=x),\ (x)(y)(x+y'=x+y'),\ (x)(x\cdot 0=0),\ (x)(y)(x\cdot y'=xy+x)$, in slightly abbreviated notation]. The "axiom" of mathematical induction says:

$$\text{(II)} \qquad S_0\cdot(x)(y)(S_x\cdot(y=x').\rightarrow S_y):\rightarrow.(x)S_x$$

where S_x is any wff not containing 'y', S_y contains 'y' where and only where S_x contains free 'x', and S_0 contains the individual constant '0' wherever S_x contains free 'x'. Thus (II) "says": "if 0 satisfies the formula S_x and, for every x, when x satisfies S_x so does $y=x+1$, then every number satisfies S_x."

Although Peano would have considered this a single "axiom," to write it down we have to write down an infinite set of wffs—one instance of (II) for each wff S_x that can be built up out of the symbols 0,E,S,T,F and logical symbols. Thus Peano arithmetic has an infinite set of axioms (and it has been proved that no finite alternative set of axioms exists). However, the usual set of axioms is *recursive*. To decide whether or not a wff is an axiom we see if it is one of the axioms that are not of the form (II) (there are only seven of these to check), and, if it is not, we then see whether or not the wff in question has the form (II) (which can be effectively decided). Thus theories with an infinite set of axioms play an important role in actual mathematics; however, it is always required in practice that the set of axioms be recursive. For, if there were no procedure for telling whether or not a wff was an axiom, then we could not tell whether or not an arbitrary sequence of wffs was a proof!

(4) A set is *recursively enumerable* if the members of the set are also the elements S_1, S_2, S_3, . . . of some sequence that can be effectively produced (i.e., produced by a Turing machine that is allowed to go on "spinning out" the sequence forever). For example, the numbers, such as 159, whose digits occur successively in the decimal expansion of $\pi = 3.14159\ldots$ are a recursively enumerable set, and they can be arranged in the sequence 3, 1, 31, 14, 4, 314, 141, 5, 41, 3141, . . . It is not known whether or not this set of integers is recursive, however. In fact, no one knows whether 7777 occurs in the decimal expansion of π or not.

The set of theorems of T, where T is a finitely axiomatized theory, is also a recursively enumerable set, and the theorems can be arranged in the sequence: (axioms of T, in lexicographic order), (theorems that can be obtained by one application of a rule of inference), (theorems that can be obtained by two applications of a rule of inference), . . . (If rules of inference such as the

"TF" of Quine's *Methods of Logic* are permitted, which can lead to infinitely many different results from a single finite set of wffs, then at the nth stage we write down only formulas of length less than 10^n which satisfy the condition given).

The set of theorems of T, where T is any *recursively* axiomatized theory, is also a recursively enumerable set. The idea of the proof is to arrange all the *proofs* in T in an effectively produced sequence (say, in order of increasing number of symbols). If one replaces the ith proof in the resulting sequence Proof_1, Proof_2, . . . by the wff that is proved by Proof_i, one obtains a listing of all and only the theorems of T (with infinitely many repetitions, of course—however, these can be deleted if one wishes).

Is every recursively enumerable set recursive? According to a fundamental theorem of recursive-function theory, the answer is "no." There is a recursively enumerable set D of positive integers that is not recursive. In other words, there is a sequence a_1, a_2, . . . of numbers that can be effectively continued as long as we wish, but such that there is *no method in principle* that will always tell whether or not an arbitrary integer eventually occurs in the sequence.

The set of theorems of quantification theory (first-order logic) is another example of a recursively enumerable nonrecursive set. The theorems can be effectively produced in a single infinite sequence; but there does not exist *in principle* an algorithm by means of which one can tell in a finite number of steps whether a wff will or will not eventually occur in the sequence—i.e., the "Decision Problem" for pure logic is not solvable.

We now see what the observation (I) comes to. It says that all recursively enumerable theories can be recursively axiomatized. If the theory T is recursively enumerable (this is equivalent to having a recursively enumerable set of axioms), then a *recursive* set S can be found which is a set of axioms for T.

Craig's Proof of (I). Craig's proof of (I) is so remarkably simple that we shall give it in full. Let T be a theory with a recursively enumerable set S of axioms, and let an effectively produced sequence consisting of these axioms be S_1, S_2, . . . We shall construct a new set S' which is an alternative set of axioms for T. Namely, for each positive integer i, S' contains the wff $S_i \cdot (S_i \cdot (\cdots))$, with i conjuncts S_i.

Clearly, each S_i can be deduced from the corresponding axiom in S' by the rule $A \cdot B$ *implies* A. Also, each axiom is S' can be deduced from the corresponding S_i by repeated use of the rules: A

implies $A \cdot A$, and A, B imply $A \cdot B$. It remains to show that S' is recursive.

Let A be a wff, and consider the problem of deciding whether or not A belongs to S'. Clearly, if $A \neq S_1$ and A is not of the form $(B \cdot (B \cdot \cdots))$, A is *not* in S'. If A is of the form $(B \cdot (B \cdot \cdots))$ with k Bs, then A belongs to S' if and only if $B = S_k$. So we just continue the sequence $S_1, S_2, \ldots$ until we get to S_k and compare B with S_k. If $B = S_k$, A is in S'; otherwise A is not in S'. The proof is complete!

Notice that, although we have given a method for deciding whether or not a wff A is in S', we still have no method for deciding whether or not an arbitrary wff C is in S, even though S and S' are trivially equivalent sets of sentences, logically speaking. For we don't know how long an initial segment $S_1, \ldots, S_k$ we must produce before we can say that if C is not in the segment it is not in S at all. The fact that S' is decidable even if S is not constitutes an extremely instructive example.

"Craig's theorem" can now be stated: (III) Let T be a recursively enumerable theory, and consider any division of the predicate letters of T into two disjoint sets, say $V_A = T_1, T_2, \ldots$ and $V_B = O_1, O_2, \ldots$ Let T_B consist of those theorems of T which contain only predicate letters from V_B. Then T_B is a recursively axiomatizable theory. Proof: Let $S_1, S_2, \ldots$ be an effectively produced sequence consisting of the theorems of T. By leaving out all wffs which are not in the subvocabulary V_B, we obtain the members of T_B, say as $V_1, V_2, \ldots$ Thus T_B is a recursively enumerable theory, and possesses a recursively enumerable axiomatization (take T_B itself as the set S of axioms). Then, by observation (I), T_B is recursively axiomatizable. Q.E.D.

The reader will observe that the proof assumes that the sets of predicate letters V_A and V_B are themselves recursive; strictly speaking we should have stated this. In practice these sets are usually finite sets, and thus trivially recursive.

The Alleged Philosophical Significance of Craig's Theorem. Imagine the entire "language of science" to be formalized, and divide the primitive predicates into two classes:[2] the so-called "theoretical terms," $T_1, T_2, \ldots$, and the so-called "observation

[2] A criticism of this dichotomy may be found in our paper "What Theories Are Not," in E. Nagel, P. Suppes, and A. Tarski, eds., *Logic, Methodology, and the Philosophy of Science* (Stanford, Calif.: Stanford Univ. Press, 1962), pp. 240–251; reviewed in this JOURNAL, **61**, 2 (Jan. 16, 1964): 79–88.

terms," $O_1, O_2, \ldots$. Let us also imagine the *assertions* of science to be formalized in this language, so as to comprise a single theory T. T would, of course, be a very big theory, including everything from psychology to paleontology and from quantum mechanics to galactic astronomy. The "predictions" of T are, presumably, to be found among those theorems of T which are in the vocabulary $V_O = O_1, O_2, \ldots$.

Let T_O be the subtheory consisting of all those theorems of T which are expressible in the observational vocabulary V_O. Thus a statement belongs to T_O just in case it meets two conditions: (1) the statement must contain no theoretical terms; (2) the statement must be a consequence of the axioms of T. Then "Craig's theorem" asserts that T_O is itself a recursively axiomatizable theory—and clearly T_O contains all the "predictions" that T does!

This had led some authors to advance the argument that, since the purpose of science is successful prediction, *theoretical terms are in principle unnecessary*. For, the argument runs, we could (in principle) dispense with T altogether and just rely on T_O, since T_O implies all the predictions that T does. And T_O contains no theoretical terms.

Thus, Hempel, discussing Craig's method,[3] writes: "Craig's result shows that no matter how we select from the total vocabulary V_T of an interpreted theory T a subset V_B of experiential or observational terms, the balance of V_T, constituting the 'theoretical terms,' can always be avoided in sense (c)" (699). This sense Hempel calls "functional replaceability" and defines as follows: "The terms of T might be said to be avoidable if there exists another theory T_B couched in terms of V_B which is 'functionally equivalent' to T in the sense of establishing exactly the same deductive connections between V_B sentences as T" (696–697).

It must be emphasized that Hempel does not rest content with this conclusion. He advances the argument we are considering only in order to reply to it. Hempel's objections are twofold: (1) the axioms of T_O (i.e., Hempel's "T_B") are infinite in number, although effectively specified, and are "practically unmanageable." (2) The *inductive* connections among sentences of T_B may be altered when T_B is considered in the light of T. Purely observational sentences may bear probabilistic confirmation relations to one another by virtue of a theory T containing "theoretical terms" which they would hardly bear to one another if they were con-

[3] C. G. Hempel, "Implications of Carnap's Work for the Philosophy of Science," in P. A. Schilpp, ed., *The Philosophy of Rudolf Carnap* (La Salle, Ill.: Open Court, 1963), pp. 685–707.

sidered simply in themselves, by someone ignorant of T. And science, Hempel argues, aims at systematizing observational sentences not only deductively but also *inductively*.

Speaking to Hempel's first objection, Israel Scheffler [4] writes:

> It seems to me, however, that if we take these programs [i.e., "specific empiricist programs"—H. P.] as requiring simply the reflection of non-transcendental assertions into replacing systems without transcendental terms, then we do not distort traditional notions of empiricism, and we have to acknowledge that Craig's result does the trick; . . . further cited problems remain but they are independent of empiricism as above formulated (170).

The discussion of Craig's theorem in Nagel's *Structure of Science* [5] rests upon the misconception that "the axioms A^* of L^* [our T_O —H. P.] . . . are specified by an effective procedure . . . upon the true observation statements W_O of L" (136). It is not the totality of *true* sentences of L in the vocabulary V_O that has to be generated as an effectively produced sequence to obtain the recursive axiomatization of T_O, but only the totality of sentences of L in the vocabulary V_O that are *theorems of* T. Thus Nagel's criticism: "Moreover, in order to specify the axioms for L^* we would have to know, *in advance* of any deductions made from them, *all* the true statements of L^*—in other words, Craig's method shows us how to construct the language L^* only *after* every possible inquiry into the subject matter of L^* has been completed" (137), is simply incorrect. We do not have to know "all the true statements" in the vocabulary V_O in order to construct a T_O with the "same empirical content" as T—indeed, we need only be given the theory T. Craig's method can be applied whether the predictions of T are in fact true or false, and is a purely formal method—thus it is not necessary to complete *any* "inquiry into the subject matter of L^*" in order to "construct the language L^*."

The Argument Reconsidered. To us there is something curiously unsatisfactory about the entire body of argumentation just reviewed. What all the participants in the debate appear to accept is the major premise (inherited from Mach and Comte, perhaps) that "the aim of science" is successful prediction or deductive and inductive systematization of observation sentences

[4] Israel Scheffler, "Theoretical Terms and a Modest Empiricism," in A. Danto and S. Morgenbesser, eds., *Philosophy of Science* (Cleveland, Ohio: Meridian Books, 1960), pp. 159–173. See also "Prospects of a Modest Empiricism," *Review of Metaphysics*, **10**, 3 (March, 1957): 383–400, and **10**, 4 (June, 1957): 602–625.

[5] New York: Harcourt Brace, 1961; reviewed in this JOURNAL, **59**, 8 (Apr. 12, 1962): 216–223.

or something of that kind. Given this orientation, the only possible reply to the fantastic proposal to employ T_O and scrap T (thereby leaving out of science all references to such "theoretical entities" as viruses, radio stars, elementary particles, and unconscious drives, to mention only a small sample) is the kind of thing we have just cited—that T_O is "unmanageable" or not "heuristically fruitful and suggestive"—or that T_O does not in fact lead to the same predictions as T when probability (confirmation) relations are considered as well as absolute (deductively implied) predictions. But surely the most important thing has not been mentioned—that leaving out of science all mention of what cannot be seen with the naked eye would be leaving out just about *all* of science.

Let us spell this out a little bit. The use of such expressions as 'the aim of science', 'the function of scientific theories', 'the purpose of the terms and principles of a theory', is already extremely apt to be misleading. For there is no *one* "aim of science," no one "function of scientific theories," no one "purpose of the terms and general principles" of a scientific theory. Different scientists have different purposes. Some scientists are, indeed, primarily interested in prediction and control of human experience; but most scientists are interested in such objects as viruses and radio stars in their own right. Describing the behavior of viruses, radio stars, etc., may not be THE "aim of science," but it is certainly an aim of scientists. And in terms of this aim one can give a very short answer to the question, Why theoretical terms? Why such terms as 'radio star', 'virus', and 'elementary particle'? *Because* without such terms we could not speak of radio stars, viruses, and elementary particles, for example—and we *wish* to speak of them, to learn more about them, to explain their behavior and properties better.

Is the "Short Answer" Too Short? We fear that a great many philosophers of science would regard our "short answer" to the question just posed as *too* short for two reasons: (1) it presupposes the existence of theoretical entities; and (2) it presupposes the intelligibility of theoretical terms. The second reason is the crucial one. Hardly any philosopher who felt no doubts about the intelligibility of such notions as "radio star," "virus," and "elementary particle," would question the existence of these things. For surely the existence of radio stars, viruses, and elementary particles is well established; indeed, one can even *observe* viruses, with the electron microscope, observe radio stars, by means of radio telescopes, observe the tracks left by elementary particles

in a cloud- or bubble-chamber. None of this amounts to *deductive proof,* to be sure; but this is an *uninteresting* objection (those of us who accept the scientific method are not unhappy because we cannot prove deductively what Hume showed to be incapable of deductive proof). Don't we have just about the best possible reasons for believing in the existence of radio stars, viruses, and elementary particles? If we don't, what would be *better* reasons?

This is not a cogent line to take, however, if the position is that no "reasons" *could be* any good because we can't believe what we can't *understand,* and so-called "theoretical discourse" is really unintelligible. That theoretical terms may be unintelligible is suggested by Scheffler's use of the expression 'transcendental term' (as if radio stars first made their appearance in Kantian metaphysics!), and his discussion shows that he takes this "possibility" quite seriously:

> To what extent is the pragmatist position in favor of a broader notion of significance positively supported by the arguments it presents? Its strong point is obviously its congruence with the *de facto* scientific use of transcendental theories and with the interdependence of parts of a scientific system undergoing test. These facts are, however, not in themselves *conclusive* evidence for significance, inasmuch as many kinds of things are used in science with no implication of cognitive significance, i.e., truth-or-falsity; and many things are interdependent under scientific test without our feeling that they are therefore included in the cognitive system of our assertions. Clearly "is useful," "is fruitful," 'is subject to modification under test," etc., are applicable also to nonlinguistic entities, e.g., telescopes and electronic computers. On the other hand, even linguistic units judged useful and controllable via empirical test may conceivably be construed as non-scientific machinery, and such construction is not definitely ruled out by pragmatist arguments (*op. cit.*, 163).

If it is really possible, however, that the term 'virus' is *meaningless* ("nonsignificant"), all appearance to the contrary notwithstanding, why is it not also possible that observation terms, such as 'chair' and 'red' are really meaningless? Indeed, traditional empiricists sometimes suggested that even such terms ("observation terms in thing language") might also be "nonsignificant," and that the only *really* significant terms might be such terms as 'pain', referring to sensations and states of mind. The traditional empiricist reason for regarding it as beyond question that *these* terms, at least, are "significant" is that (allegedly) in the case of these terms the meaning is the referent, and I know the referent from my own case. However, this argument is self-undercutting: if 'pain' means what *I* have, then (on the traditional view) it is logically possible that no one else means by 'pain' what I do. Indeed, there may be *no* such thing as the

public meaning of the word 'pain' at all! We believe that this traditional theory of meaning is in fact bankrupt and that it cannot account for the meaning of *any* term, least of all for sensation words.

In point of fact, a term is "meaningful"—i.e., has meaning in the language—if it belongs to the common language or has been explained by means of terms already in the common language. The fact that we can and do have theoretical terms in our language rests upon the fact that there was never a "pretheoretical" stage of language; the possibility of talking about unobservables is present in language from the beginning. 'Radio star', 'virus', and 'elementary particle' are perfectly good (meaningful) expressions in English. If you have doubts, consult your nearest dictionary! "But maybe these words are *really* meaningless." Well, they are defined in the dictionary. "Maybe the definitions are meaningless too." What sort of a doubt is this?

We are not urging that every word in the dictionary has a place in science or that the concept of "meaning" in linguistic theory does not require further investigation (cf. Paul Ziff, *Semantic Analysis*,[6] for a pioneer investigation). But we are urging that the concept of meaning used in everyday life and in linguistic theory is also the one appropriate to philosophy and that every term that has an established place in English syntax and an established use has meaning—that's how we *use* the word 'meaning'. The word 'God', for example, certainly has meaning. It does not follow that it belongs in science—but it does not belong *not* because it is "nonsignificant," but just for the obvious reason—that there are no scientific procedures for determining whether or not God exists or what properties He has if He does exist. Whether or not the word *should* be employed as it is in *non*scientific contexts is another question. However, in the case of 'radio star' one could hardly argue "there are no scientific procedures for detecting radio stars." Thus 'radio star' is not only an expression with a meaning, but one which is usable in science. To suggest that it might be unintelligible is to promulgate an *unintelligible concept of intelligibility*.

Scheffler tries to make clear the notion of significance he has in mind by writing "cognitive significance, i.e., *truth-or-falsity*," but this is no help at all. For " 'Radio stars exist' is neither true nor false" means " 'Radio star' is not a (sufficiently) clear expression" if it means anything. And the reply to a charge that

[6] Ithaca, N.Y.: Cornell Univ. Press, 1960; reviewed in this JOURNAL, **59**, 1 (Jan. 4, 1962): 5–20.

'radio star' was unclear would surely be, "In what *way* is it unclear?" Theoretical terms can't just be "unclear"; there has to be some relevant respect in which they might be sharpened and have not been.

We conclude that (a) theoretical terms are intelligible, in any ordinary sense of that term; and (b) theoretical entities have been established to exist (up to a sophomorish kind of skeptical doubt). Thus our "short answer" stands: theoretical terms are not eliminable, "Craig's theorem" notwithstanding, if one wishes to talk about theoretical entities; and we *do* wish to talk about theoretical entities.

HILARY PUTNAM

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

SALMON'S VINDICATION OF INDUCTION

THE conclusions urged in Wesley Salmon's recent work on induction are so striking that it may be worth recording some difficulties. His claim * to rescue Reichenbach's vindication of induction has met with some skepticism,[1] but objections to it have been of an entirely philosophical nature; many, for example hinge on Goodman's riddles about 'grue'. It may not have been noticed that Salmon's argument is defective for much less sophisticated reasons: its invalidity can be shown without doing any philosophy at all. Several recent writers [2] seem to imply that, aside from Goodman's riddles, Salmon has at least proved that Reichenbach's straight rule for estimating long-run frequencies is preferable to any rival rule. This idea had better be scotched before it creeps into the textbooks.

The Straight Rule. An *estimator* of some kind of magnitude (like length or long-run frequency) is a rule for making estimates

* In "Vindication of Induction," in H. Feigl and G. Maxwell, eds., *Current Issues in the Philosophy of Science* (New York: Holt, Rinehart and Winston, 1961); and "On Vindicating Induction," in H. E. Kyburg and E. Nagel, eds., *Induction: Some Current Issues* (Middletown, Conn.: Wesleyan Univ. Press, 1963).

[1] S. Barker, "Comments on Salmon's 'Vindication of Induction'," in Feigl and Maxwell; and discussion of Salmon's paper by other speakers in Kyburg and Nagel.

[2] H. E. Kyburg, Jr., "Recent Work in Inductive Logic," *American Philosophical Quarterly*, **1**, 4 (October, 1964): 249–287; and F. Schick, review of J. J. Katz, *The Problem of Induction and Its Solution*, this JOURNAL, **60**, 16 (Aug. 1, 1963): 473–478.