ON THE USE OF HILBERT'S ε-OPERATOR IN SCIENTIFIC THEORIES

RUDOLF CARNAP

Los Angeles, California

1. The ε-operator in logic and set theory

Hilbert's ε-operator is usually applied only in logic and in mathematical theories, especially arithmetic and set theory. The first part of this paper makes some comments on its use in these fields. In the second part I shall point out the possibility of a useful application of the ε-operator in the formulation of theories in empirical science.

Hilbert [14], [15] introduced the ε-operator in such a way that, if φ is any satisfiable open sentential formula with 'x' as the only free variable, then the ε-term denotes '$\varepsilon_x \varphi$' denotes an object that satisfies φ. This object may be regarded as the representative of φ, or of the corresponding set if there is such a set.

The following two formulas may be taken as logical *axioms* for the ε-operator:

(1) $(\exists y)Fy \supset F(\varepsilon_x Fx)$,

(2) $(x)(Fx \equiv Gx) \supset \varepsilon_x Fx = \varepsilon_x Gx$.

Axiom (1) is Hilbert's axiom. Axiom (2), first proposed by Ackermann [1], says that coextensive formulas have the same representative; it is therefore a kind of extensionality axiom. Asser [2] gives the first systematic treatment of the ε-calculus, its syntactical rules, its semantical (set-theoretic) interpretation in terms of logical selection functions (a function of this kind assigns to each non-empty subset of the domain of individuals an element of this subset); and he gives simplified proofs for the most important metatheorems.

Hilbert has shown that both the existential and the universal quantifiers can be defined with the help of the ε-operator. The definition of the former is as follows:

(3) $(\exists x)Fx \equiv F(\varepsilon_x Fx)$

Let PK be the ordinary predicate calculus with sentential connectives, the two quantifiers, and the sign of identity as primitive logical constants. Let the ε-calculus PK_ε contain, in addition, the ε-operator as primitive, and the additional axioms (1) and (2). Hilbert and Bernays [16] have pointed out that the system of rules of PK_ε is simpler than that of PK. They also showed the great usefulness of the ε-operator for metamathematical investigations.

Axiom systems of *set theory* are usually constructed on the basis of PK, using a language L formed from the language of PK by adding the primitive constant 'ϵ' as the sign of the membership relation. As an alternative, PK_ε may be used, and the language L_ε formed from the language of PK by the addition of 'ε'.

Since the ε-operator expresses an (unspecified) selection function, there is clearly a close connection between this operator and the axiom of choice. Some who regard this axiom with suspicion are therefore also reluctant to accept the ε-operator. Fraenkel has long ago emphasized the "purely existential character" of the axiom of choice ([10]; [11, p. 54 ff.]): it asserts merely that, if the set x fulfills certain conditions, then there is at least one selection set of x, without asserting that for every such set x it is possible to specify a particular selection set of x. Analogously, the ε-expression for a satisfiable formula refers to a representative without specifying it. It is indeed important to recognize the special character of the axiom of choice in contrast to other axioms, and likewise the special character of the ε-expressions, their indeterminacy, which we shall discuss later. But these features are in themselves not sufficient reasons for rejecting either the axiom of choice or the ε-operator.

What now is the connection between the ε-operator and the axiom of choice? Is the acceptance of the former tantamount to that of the latter? In more formal terms, is the axiom of choice derivable from the other axioms of set theory if the underlying logic contains the ε--operator with its axioms? In some sense this is the case, but the assertion needs some qualifications. Fraenkel and Bar-Hillel, in an illuminating brief exposition of the ε-operator, are certainly right with the following statement [11, p. 184]: "There is no reason to suppose that in a set theory constructed on the basis of an ε-calculus the principle of choice would become generally derivable, unless the specific axioms of that set theory contain ε-terms themselves." The decisive point for this question of derivability is the specific form of the axiom schema of subsets (*Aussonderungsaxiom*). In the customary language L it may be formulated as follows, where "Su" stands for "u is a set":

(4) '$Su \supset (\exists y)[Sy \bullet (v)(v \in y \equiv v \in u \bullet \varphi)]$' where φ is any sentential formula of language L containing 'v' as the only free variable.

If L_ε is taken as the axiomatic language, there is the choice of two versions of the axiom schema, differing in the kinds of formulas admitted as φ. The first version is the same as (4): only formulas of L_ε without 'ε' are admitted; in other words, formulas of L (as a sub-language of L_ε). The second version, which we shall call. (4_ε), is formed from (4) by replacing 'L' with 'L_ε'. (4_ε) is stronger than (4). But to accept this version seems natural, once the ε-operator has been accepted as a primitive logical constant.

Consider now the principle of choice:

(5) If x is a set such that:
(a) any element of x is non-empty,
(b) any two distinct elements of x are disjoint,
then there is a set y (called a selection set of x) such that
(c) $y \subset \bigcup x$,
(d) for any element z of x, $y \cap z$ has exactly one element.

It can now be seen easily that, if the axiom schema of subsets is taken in the stronger form (4_ε), then (5) is derivable. The derivation is as follows. Let x be any set satisfying the conditions (a) and (b) in (5). According to the axiom of the union set, $\bigcup x$ is a set. Therefore, by (4_ε), there is a set y containing exactly those elements v of $\bigcup x$ for which

$$(\exists z)\ [z \in x \bullet v = \varepsilon_u(u \in z)].$$

(This last formula is taken as φ in (4_ε).) Thus y is a subset of $\bigcup x$ containing just the representatives of the elements of x. Hence y satisfies the conditions (c) and (d) in (5). Thus (5) is derived.

2. The ε-operator in a scientific theory

Let L be a language suitable for the formulation of a given scientific theory. We divide the primitive descriptive terms of L in the customary way (compare [5],[12]) into observational terms (*O*-terms, e.g. "hot", "blue") and theoretical terms (*T*-terms, e.g. "temperature", "electric field”). The interpretation of *O*-terms is assumed to be known. If their meanings are not logically independent, it is assumed that their meaning relations are expressed by *A*-postulates (analytic postulates or meaning postulates, [3]). Let A_O be the conjunctions of the *A*-postulates for the *O*-terms. An *O*-sentence (i.e., one containing only *O*-terms, no *T*-terms) is said to be *A*-true (analytically true) if it follows logically from A_O.

The *T*-terms of a scientific theory are introduced by the postulates of the theory. These postulates are of two kinds, the theoretical postulates (*T*-postulates, containing only *T*-terms, no *O*-terms) and the correspondence postulates (*C*-postulates, containing both *T*-terms and *O*-terms). Let *T* be the conjunction of the *T*-postulates of a given theory, *C* that of its *C*-postulates, and *TC* that of both kinds of postulates. Thus the first form of the total set of postulates is this:

FORM I: A_O, T, C.

In addition, there are axioms and rules, not to be stated here, of an elementary logic (e.g. *PK*) and a higher logic, either in set-theoretic or in type-theoretic form, sufficient for the scientific theory in question.

Suppose that *TC* contains *n* *T*-terms, say ‘t_1’,...,‘t_n’, and *m* *O*-terms, say ‘o_1’,..., ‘o_m’. Then *TC* can be written as follows:

(6) TC: $\Phi(t_1,...,t_n; o_1....,o_m)$,

where ‘Φ’ is a purely logical $(n + m)$-place predicate. F. P. Ramsey [17] has proposed to use instead of *TC* the following sentence *R*, which we shall call the Ramsey sentence of the theory:

(7) R: $(\exists u_1)...(\exists u_n)\ \Phi\ (u_1,...,u_n; o_1,...,o_m)$,

where ‘u_1’,..., ‘u_n’ are variables corresponding to ‘t_1’,..., ‘t_n’, respectively. Ramsey has shown that every *O*-sentence which follows logically from *TC*, follows also from the weaker sentence *R*, which is itself an *O*-sentence. The methodological properties of the Ramsey sentence are discussed in detail by Hempel [12] [13].

The *T*-terms have no direct interpretation since they refer to unobservables. What interpretation they possess is given to them by the *C*- and *T*-postulates. And this interpretation is incomplete, because the scientist can always add further *C*-postulates (e.g., operational rules for T-terms) or T-postulates and thereby increase the specification of the meanings of the T-terms. In view of this situation, some authors have expressed doubts about the possibility of extending the concept of A-truth to sentences with T-terms. I have proposed ([6],[8, §24] which is my reply to Hempel [13]) to take as the only A-postulate A_T for T-terms the conditional with R and TC as components:

(8) A_T is $R \supset TC$:

$$(\exists u_1)\ldots(\exists u_n)\ \Phi\ (u_1,\ldots,u_n;\ o_1,\ldots,o_m) \supset \Phi\ (t_1,\ldots,t_n;\ o_1,\ldots,\ o_m).$$

With the help of this postulate A_T I define:

(9) Let S be any sentence (with or without T-terms) of language L.
S is A-true in $L =_{Df} S$ follows logically from A_O and A_T.

Thus we obtain a new form of our postulate system:

FORM II: A_O, R, A_T.

It is easily seen, that R and A_T together are logically equivalent to TC. R has factual content, but does not contribute to the interpretation of the T-terms; conversely, A_T does the latter but not the former.

For our further discussion it is convenient to define 't' and 'o', as follows:

(10) $t = <t_1,\ldots,\ t_n>$.
(11) $o = <o_1,\ldots,\ o_m>$.

We shall use 'u' as a variable corresponding to 't'. Then (6), (7) and (8) are formulated as follows (for simplicity, we use here for the two-place predicate the same symbol 'Φ' as for the original ($n + m$)-place predicate):

(12) TC: $\Phi(t,o)$.

(13) R: $(\exists u)\ \Phi\ (u,o)$.

(14) A_T: $(\exists u)\ \Phi\ (u,o) \supset \Phi(t,o)$.

I shall now give a new form III of the system, which uses the language L_ε. It will be possible, with the help of the ε-operator, to give n + 1 definitions $A^0_T, A^1_T, \ldots, A^n_T$ for 't', 't_1',..., 't_n', respectively. In form III, these definitions will be taken as *A*-postulates for the defined terms just mentioned. (It seems generally to have some advantages to regard explicit definitions of descriptive terms as a special kind of *A*-postulates.) The conjunction A'_T of these n + 1 *A*-postulates takes in form III the place of A_T in the earlier form II, while A_O and *R* remain unchanged:

FORM III: A_O, R, A'_T.

According to (14), A_T says this: "If there is anything that bears the relation Φ to *o*, then *t* does so." This, in view of (1), suggests defining *t* as the ε-representative of Φ (.,*o*):

(15) A^O_T: $t = \varepsilon_u\, \Phi\, (u,o)$.

The logical predicate 'Φ' is supposed to be defined in such a way that its first-place arguments are *n*-tuples. Hence by (15) *t* is an *n*-tuple. We now define t_i ($i = 1,\ldots,n$) as the *i*-th member of the *n*-tuple *t*. Therefore we take as the definitions $A^1_T, \ldots, A^n_T$ the instances of the following schema, with $i = 1,\ldots,n$:

(16) A^i_T: $t_i = \varepsilon_x\, [(\exists u_1)\ldots(\exists u_n)\, (t = <u_1,\ldots,u_n> \bullet x = u_i)]$.

(Instead of the operator 'ε_x' we could use here the customary description operator '(ι*x*)', since the formula in square brackets fulfills the uniqueness condition with respect to '*x*'.)

Since *t* is an *n*-tuple, we obtain from the *n* definitions of the schema

(16):

(17) $t = <t_1,\ldots,t_n>$

This corresponds to the earlier definition (10) in *L*. Here in form III, we cannot use the definition (10), because 't_1', etc. are now defined on the basis of '*t*'. However, we still use here the definition (11).

It will now be shown that the earlier sentence A_T can be derived from the new A -postulates A'_T, i.e., (15) and (16). Thus this sentence, which was an A-postulate in form II, is still A-true in the present form III, even though it is no longer an A-postulate.

We have by the first ε-axiom (1):

(18) $(\exists u)\ \Phi\ (u,o) \supset \Phi\ (\varepsilon_u\ \Phi\ (u,o),o)$.

(Hence with (15):

(19) $(\exists u)\ \Phi\ (u,o) \supset \Phi\ (t,o)$,

which is A_T in the form (14). (We can also obtain A_T in the original form (8), with the help of (17) and (11).)

Thus from A'_T alone we have derived A_T, which is $R \supset TC$. By using now the synthetic postulate R, which occurs also in form III, we obtain TC. This shows that the postulates of form III are sufficient to recover all T- and C-postulates of the original form I.

In form III we can clearly recognize that the sentences chosen as A-postulates for 't' and the T-terms do not contribute anything to the factual content of the theory but serve merely for the (partial) specification of meanings, since these sentences have the form of definitions.

It may at first glance seem surprising that it should be possible to give explicit definitions of the T-terms. This might appear incompatible with the fact that the interpretation of the T-terms is purposively left incomplete. The frequently made assertion that it is impossible to define the T-terms is indeed correct *if* the primitive logical constants of the language used have a well determined complete interpretation, as is usually the case. This intended interpretation may either be given formally by semantical rules formulated in the metalanguage or by informal explanations. Let us call a constant "determinate" (with respect to interpretation) if it has such an interpretation of either kind, and otherwise "indeterminate".

If a language like L_ε is used, the situation is radically different, because the symbol 'ε' was intentionally introduced by Hilbert as an indeterminate constant. Its meaning is specified by the axioms (1) and (2) only to the extent that any non-empty set has exactly one representative and that this representative is an element of the set. If the set has more than one element, then nothing is said, either officially or unofficially, as to which of the elements is the representative. Thus, for example, $\varepsilon_x(x = 1 \vee x = 2 \vee x = 3)$ must be either 1, or 2, or 3; but there is no way of finding out which it is.

Indeterminate logical constants are not often used. One other example is the iota-operator for individual descriptions, provided the rules for it are such that the following two conditions are fulfilled: (1) all descriptions whose operands do not satisfy the condition of uniqueness denote one and the same individual (a method proposed by Frege, in contrast to that of Russell), and (2) this individual is not specified (in contrast to Frege's method). One of the methods for descriptions which I have proposed is of this kind ([4, p. 37]; [7, §35], *but not in the first edition or in the English edition*). In this one of my methods (in distinction to others in which the individual in question is specified, e.g., as 0 or as a_o) the description '$(\iota x)(x \neq x)$' and the constant 'a^*' *L*-equivalent to it are indeterminate.

The postulates *TC* are intended by the scientist who constructs the system to specify the meaning of '*t*' to just this extent: if there is an entity satisfying the postulates, then '*t*' is to be understood as denoting one such entity. Therefore the definition (15) gives to the indeterminate constant '*t*' just the intended meaning with just the intended degree of indeterminacy.

BIBLIOGRAPHY

[1] ACKERMANN, W., 1924-25, Die Widerspruchsfreiheit des Auswahlaxioms, *Nachrichten d. Ges. d. Wiss. zu Göttingen,* math.-phys. K1., 246-250.

[2] ASSER, G., 1957, Theorie der logischen Auswahlfunktionen, *Zeit. f. math. Logik u. Grundl. d. Math*., 3, 30-68.

[3] CARNAP, R., 1952, Meaning postulates, *Phil. Studies*, 3, 65-73. Repr. in [4].

[4] CARNAP, R., 1956, *Meaning and necessity,* 2nd ed., Univ. of Chicago Press, Chicago.

[5] CARNAP, R., 1956, The methodological character of theoretical concepts. In [9].

[6] CARNAP, R., 1958, Beobachtungssprache and theoretische Sprache, *Dialectica,* **12**, 236-248. Also in: 1959, *Logica, Studia Paul Bernays dedicata* (Bibliothèque Scientifique, vol. 34), Neuchâtel.

[7] CARNAP, R., 1960, *Einführung in die symbolische Logik,* 2nd ed., Springer,Wien.

[8] CARNAP, R., 1961, Replies and systematic expositions. Forthcoming in [18].

[9] FEIGL, H., et al. (eds.), *Minnesota Studies in the Philosophy of Science,* 1956, vol. I; 1958, vol. II, Minneapolis.

[10] FRAENKEL, A. A., 1928, *Einleitung in die Mengenlehre,* 3rd ed., Springer, Berlin (Repr. New York, 1946).

[11] FRAENKEL, A. A. AND BAR-HILLEL, Y., 1958, *Foundations of Set Theory,* North--Holland Publishing Co., Amsterdam.

[12] HEMPEL, C. G., 1958, The theoretician's dilemma. In [9], Vol. II.

[13] HEMPEL, C. G., 1961, Implications of Carnap's work for the philosophy of science. Forthcoming in [18].

[14] HILBERT, D., 1923, Die logischen Grundlagen der Mathematik, *Math. Annalen,* **88**, 151-165.

[15] HILBERT, D., 1928, Die Grundlagen der Mathematik, *Abhandl. math. Seminar d. Hamburg. Universität,* **6**, 65-92. Repr. as: *Hamburger math. Einzelschriften,* No. 5, Hamburg, 1928.

[16] HILBERT, D. AND BERNAYS, P.,1939, *Grundlagen der Mathematik,* Vol.II, Springer, Berlin (Repr. *Ann Arbor, Mich.,* 1944).

[17] RAMSEY, F. P., 1929, Theories, Posthum. publ. as ch. IX A in: *The foundations of mathematics,* Kegan, Trench, Trubner & Co., London, 1931.

[18] SCHILPP, P. A. (ed.), *The Philosophy of Rudolf Carnap* (The Library of Living Philosophers). Forthcoming.