by the same author

THE PHILOSOPHY OF RELATIVITY

THE PROBLEMS OF LOGIC

by

ANDREW PAUL USHENKO

London

GEORGE ALLEN AND UNWIN LTD

re-examination. Also logical intuition shows an almost unlimited ingenuity. A failure of a postulational system to deal with the paradoxes would be equivalent to a condemnation of the system. But if intuition has so far failed, we can say that we have not yet discovered the intuitive principle which is relevant to the problem.

Chapter III

CONSISTENCY AND THE DECISION-PROBLEM

§ 1. Introduction

The problem of consistency is a peculiarity of the postulational logic. When the principles of logic are accepted on purely intuitive grounds, there can be no question concerning their truth and significance. And if they are both true and significant, they cannot lead to inconsistency. Even in the semi-postulational procedure of logistic there is no need for a proof of consistency. Once the theory of types is introduced in order to insure significance, the primitive propositions of a logistic system are known to be mutually consistent because they are recognized to be true by immediate inspection. In a postulational system of logic, on the other hand, the restrictions of significance are not derived from ontological distinctions; they are incorporated into the arbitrary rules of procedure. The starting-point of the procedure is formed by the postulates which, like the rules, are also accepted by convention. Under these circumstances it is always possible to expect, unless a proof to the contrary is given, the emergence of a paradox, which was not anticipated when the conventions were introduced, or of a contradiction, which would disclose that some conventions in the basis of the system happen to clash with others. But within the system the demonstration of theorems may have gone a long way before one stumbles against inconsistency. Obviously one wants to find the weakness of the foundation not by means of the collapse of the building but before one has started to build. For this reason the question whether the postulational system is contradictory must be decided by theoretical (or metalogical) considerations and not through the procedure carried on within the system.

A proof of consistency is a demonstration that with the postulates and rules of the formal system no two theorems can be deduced which contradict one another. Thus to prove that the calculus of unanalysed propositions, i.e. the calculus whose variables, p, q, etc., are interpretable as propositions, is consistent, one must show that p and $\sim p$ are not both theorems. As explained in § 2, if both p and $\sim p$ are deducible, then any formula q is deducible or is a theorem. Conversely, if a system contains a formula which is not a theorem, the system must be consistent. The presence of unprovable formulas might be suggested as a practical criterion of consistency, provided one could always decide whether a given formula is provable or not. But the possibility of such a decision is itself a problem, known as the "*decision-problem*" (Entscheidung problem). The question is whether there exists, with regard to any given formula of a postulational system of logic, a finite procedure, i.e. a procedure which takes a finite number of steps, whereby one can determine either that the formula is a *theorem* or that it is not deducible. If the formula is a theorem, its interpretation

must give true propositions for all values of its variables; if the formula is not a theorem, it may be *self-consistent*, i.e. interpretable as true at least for some values of its variables, or else it is a *contradiction*. To establish that the formula is consistent is, of course, the same thing as showing that its contradictory is not a theorem. Now although one cannot tell off-hand whether the decision-problem is solvable, it is easy to anticipate the general conditions which a solution would have to satisfy. One condition is the existence of a property, let us call it the *K*-property, which only deducible formulas of the system have; the other condition is the possibility of establishing by a finite procedure with regard to any given formula whether it has the *K*-property. The solution of the decision-problem and of the problem of consistency does not yet mean that the postulational system is adequate as logic. Logic aims at a system of principles which are always true, i.e. it must consist of formulas which are true for all values of the variables. Hence the postulates for logic must enable one to sort without residue all the formulas within the system, i.e. all the formulas expressed in terms of the undefined elements of the system, into principles which are theorems and formulas whose claim to be always true is unfounded and which, therefore, are refutable. A system every formula of which is either deducible or refutable is called *complete*. The relation of completeness to consistency is this: completeness insures that at least one, consistency that at most one, of two contradictory formulas is a theorem.

Confusion in a discussion of consistency might easily arise if the distinction between *postulational logic* and *postulational mathematics* is not brought out. A postulational system is a logic if it contains, besides variables, constant symbols which are interpretable as the propositional connectives "if-then" and "it is false that", or some equivalents of these, and the logical properties of which are defined by some of the postulates. A system is a mathematics if its constants are interpretable as mathematical operators or relations, such as "plus", "greater than", and the like, which do not connect propositions with one another. The presentation of the postulates for a mathematics requires either the medium of a symbolic logic or of an ordinary language. For example, the postulates for serial order can be given in terms of the undefined relation "$<$", by three statements in English:

(1) Given a class of elements K, if a and b are not the same elements of K, then either $a < b$ or $b < a$.

(2) Given a class of elements K, if $a < b$, then a and b are not the same elements of K.

(3) Given that a, b, c are elements of K, if $a < b$ and $b < c$, then $a < c$.

The same system can be formulated in the symbolism of the *Principia Mathematica*, if we let "$f\,x$" symbolize "x is an element of K" and "$g\,(x, y)$" stand for "$x < y$", as follows:

(1′) $[f\,a\,.\,f\,b\,.\sim(a = b)] \supset [g\,(a, b) \vee g\,(b, a)]$.

(2′) $[f\,a\,.\,f\,b\,.\,g\,(a, b)] \supset \sim(a = b)$.

(3′) $[f\,a\,.\,f\,b\,.\,f\,c\,.\,g\,(a, b)\,.\,g\,(b, c)] \supset g\,(a, c)$.

Although the primed system is entirely symbolic, and uses the propositional connectives "$\supset$" and "$\vee$", it is not logic, because its postulates do not define the properties of these connectives. The logic of implication and disjunction is presupposed throughout.

The distinction between a postulational logic and a postulational mathematics is relevant to the solution of the problem of consistency. In general, if consistency of the linguistic or symbolic medium of presentation is assumed, consistency of postulates for a mathematics is provable. The usual method of proof is by *interpretation*. If all postulates are satisfied by some concrete example, then in virtue of the principle that actuality cannot be inconsistent, the system which it exemplifies must also be consistent. For example, the system for serial order can be exemplified by the notes within an octave, if we interpret "$<$" as "lower (in pitch) than", since it is true that:

(1) If a and b are two different notes, then either a is lower than b or b is lower than a;

(2) If a is lower than b, they are different tones;

(3) If a is lower than b and b is lower than e, then a is lower than c.

And while the notes in their serial order of pitches are actually produced any time one runs his fingers over the keyboard of a piano, the consistency of the abstract conditions for serial order is thereby demonstrated. The weakness of the method of concrete interpretation is that there can be no assurance that concrete examples, such as the notes in their order of pitches, are available for any given set of mathematical postulates. But whenever they have no concrete examples on hand mathematicians can resort to *abstract interpretation*, and this means that consistency of a postulational mathematics is always provable.*

Of course, proof by interpretation is contingent

* Abstract interpretation gives a *schema* of an example instead of concrete examples such as the order of the notes in an octave and the like. Thus with regard to the system for serial order a *schema* may be constructed by taking three items such that while a serial relation holds between one item and each of the others, taken in that order and, between one and the other of the latter items, the same relation fails to hold for any other permutation of items. To give an illustration, while the serial relation of "being lower in pitch" holds between *do* and *re*, *do* and *mi*, and *re* and *mi*, it does not hold for *do* and *do*, *re* and *do*, *re* and *re*, *mi* and *do*, *mi* and *re*, and *mi* and *mi*. This concrete illustration, however, may be taken as an abstract schema provided we use *do*, *re*, *mi* not as the names of the notes C, D, E in the key of C, but as names of any given trio of items in abstraction from their nature. The consistency of the abstract example is seen if we observe that each singular statement as to whether the relation holds or not is concerned with a different permutation of the items and, therefore, all the singular statements which form the abstract example have a different subject-matter and so cannot be inconsistent. (For a more detailed exposition and its development, cf. Paul Henle, "A Definition of Abstract Systems," *Mind*, 1935.)

upon the assumption that the logic of the medium (within which the postulates for a mathematics are introduced) is itself consistent. Thus, from the postulationalist standpoint, the basic problem of consistency is concerned with a postulational logic rather than with mathematics. But here again a distinction has to be drawn. Logic may be "*pure*", as exemplified by the calculus of unanalysed propositions and the calculus of predicate (cf. §§ 2 and 3), or combined with mathematics. The "*combined logic*" must list among its postulates some which determine the properties of certain constant symbols in a way which makes them interpretable as the basic mathematical relations of numerical equality and the like. In anticipation of the following sections it may be stated that the problems of consistency and completeness can be solved for a pure calculus, but not for a logic combined with mathematics. The proofs for "pure logic" are given in the order of increasing complexity of the systems concerned; first, for the calculus of unanalysed propositions; next, for the same calculus in combination with the logic of predicates.

§ 2. The Calculus of Unanalysed Propositions

Every symbolic postulational system is based upon a set of postulates which are expressed in terms of variables and undefined constant symbols. A set of postulates for the system of unanalysed propositions is given by the "primitive propositions" of section A of the *Principia*. In the modified version of *Grundzüge* *der theoretischen Logik* by Hilbert and Ackermann the set contains four formulas:

(*a*) $(p \vee p) \supset p$;
(*b*) $p \supset (p \vee q)$;
(*c*) $(p \vee q) \supset (q \vee p)$;
(*d*) $(p \supset q) \supset [(r \vee p) \supset (r \vee q)]$.

The primitive or undefined symbols are "$\sim$" (read "curl") and "$\vee$" (read "wedge"). The symbols "$\supset$", which appears in the formulation of the postulates, as well as the symbols "." and "$\equiv$" can always be omitted by means of the following definitions:

(1) "$p \supset q$" is defined as "$\sim p \vee q$"; (2) "$p . q$" is defined as "$\sim (\sim p \vee \sim q)$"; and (3) "$p \equiv q$" is defined as "$(p \supset q) . (q \supset p)$".

The theorems are derived from (*a*), (*b*), (*c*), (*d*) by (α) the rules of substitution and (β) the rule of inference.

(α) A given variable can be replaced at each of its occurrences within a formula by the same compound expression. For example, one can substitute "$q \vee r$" for "p" in (*a*) and derive the formula "$[(q \vee r) \vee (q \vee r)] \supset (q \vee r)$".

(β) From the conjunction of p and $p \supset q$ one can derive q.

This abstract system is called the calculus of unanalysed propositions because it is *interpretable* in terms of propositions taken as units, i.e. without regard to their constituents. Thus (*b*) can be inter-

preted as the assertion that "a proposition implies that either it or some other proposition is true".

The consistency of the calculus (taken in abstraction from its interpretations) can be proved by the test of the truth-table.* The test will convince one that (*a*), (*b*), (*c*), (*d*) are tautologies and that by means of (α) and (β) nothing but tautologies can be derived from tautologies. Hilbert and Ackermann give an analogous but simplified method of proof.

Let every variable be arbitrarily interpreted in one of two ways, as being either *0* or *1*; let "$\vee$" stand for the sign of arithmetical multiplication, and $\sim p$ be *1* if it stands for ~ 0, and *0* if it stands for ~ 1. With this interpretation every postulate gives *0*. For example, taking (*a*) in the form "$\sim (p \vee p) \vee p$", one observes that at least one side of the main wedge must be interpreted as *0*, and therefore the whole product is *0*.† Furthermore, according to

* The use of the truth-table need not depend on logical intuition provided one does not interpret its symbols *T* and *F* (or *1* and *0*) as, respectively, truth and falsehood. In abstraction from such an interpretation the truth-table is a table of permutations for two signs. The truth-function can then be defined by convention, i.e. by an arbitrary assignment of either *T* or *F* to each of their permutations. Thus we might assign *T* to the permutation *T T* and *F* to all other permutations, and define this assignment as constitutive of a conjunction.

† For, by convention, p is either *1* or *0*. Let p be *1*. Then "$\sim (p \vee p) \vee p$" is "$\sim (1 \vee 1) \vee 1$", which is "$\sim 1 \vee 1$", i.e. "$0 \vee 1$", i.e. *0*. Now let p be *0*. Then the formula is "$\sim (0 \vee 0) \vee 0$". This gives "$\sim 0 \vee 0$", i.e. "$1 \vee 0$", i.e. *0*. In either case (*a*) is *0*.

the rules (α) and (β) only *0* can be derived from the postulates each of which is *0*. Transformations by (α) cannot change either the range of arithmetical interpretation or the main structure of the original formulas, while, in applying (β), the premises p and $p \supset q$ can be identified with a pair of postulates only if each premise is *0*, and this is possible only when q is likewise *0*.* Since all theorems of the system must be interpreted as *0*, the system is *consistent*.† By a somewhat similar procedure the calculus is shown to be *complete*.

In order to solve the decision-problem one can begin by showing that every formula of the system is transformable into a standard pattern called the *conjunctive normal form*. The transformation is performed with the aid of the following rules (which are derivable from the postulates):

(a 1) The symbols "$\vee$" and "." have the associative, distributive and commutative properties;

(a 2) The symbols "$\sim(\sim p)$" and "p" can replace one another in any context;

(a 3) The symbol "$\sim(p \,.\, q)$" is replaceable by the symbol "$\sim p \vee \sim q$", and the symbol "$\sim(p \vee q)$" by "$\sim p \,.\, \sim q$";

* For if *9* were *1*, "$p \supset q$", i.e. "$\sim p \vee q$" would be "*1* $\vee$ *1*", i.e. *1*.

† If the system were inconsistent, both p and $\sim p$ would be theorems. But then at least one of them would be interpreted as *1*.

(b 1) $p \vee \sim p$ is a theorem;

(b 2) If p is a theorem and q is any formula, then $p \vee q$ is a theorem;

(b 3) If p is a theorem and q is a theorem, then $p \,.\, q$ is a theorem.

Now let an expression be given for transformation into its normal form. By means of the definitions (1) and (3) it is cleared of the symbols "$\supset$" and "$\equiv$". By means of (a 3) the sign "$\sim$" is made to precede only single variables. By means of (a 2) one gets rid of the reiterated "curl". Finally, by means of (a 1) the expression is formulated as a conjunction of disjunctions of *single proposition-variables with or without a single curl each*. This gives the normal conjunctive form. To illustrate each successive step of the transformation let the original expression be:

$$(p \supset q) \equiv (\sim q \supset \sim p).$$

The successive transformations are:

$(\sim p \vee q) \equiv (\sim\sim q \vee \sim p)$ (By def. (1)).

$(\sim p \vee q) \equiv (q \vee \sim p)$ (By (a 2)).

$[\sim(\sim p \vee q) \vee (q \vee \sim p)]$
$\quad .\,[\sim(q \vee \sim p) \vee (\sim p \vee q)]$ (By def. (1) and (2)).

$[(\sim\sim p \,.\, \sim q) \vee (q \vee \sim p)]$
$\quad .\,[(\sim q \,.\, \sim\sim p) \vee (\sim p \vee q)]$ (By (a 3)).

$[(p \,.\, \sim q) \vee (q \vee \sim p)]$
$\quad .\,[(\sim q \,.\, p) \vee (\sim p \vee q)]$ (By (a 2)).

Applications of the distributive law give the conjunctive normal form:

$$(p \vee q \vee \sim p) \cdot (\sim q \vee q \vee \sim p) \cdot (\sim q \vee \sim p \vee q) \cdot (p \vee \sim p \vee q).$$

The fact that every formula is reducible to its normal form leads to a solution of the decision problem because there exists a simple criterion which determines whether the given normal form is a theorem. It is a theorem when and only when in each member of the conjunction, i.e. in each set of disjunctions, at least one proposition-variable occurs once with and once without a curl. If this condition were not satisfied for some member of the conjunction, one could "force" this member to take the truth-value F (interpretable as "false") by assigning to each proposition-variable without a curl the value F and to each proposition-variable with a curl the value T (interpretable as "true"). And, of course, if at least one member of a conjunction is false, the whole conjunction is also false.

To give a simple example of a decision whereby the test by actual deduction is avoided let me take the formula:

$$p \supset (\sim p \supset q),$$

which is transformed into:

$$\sim p \vee p \vee q.$$

This is the normal form of the original expression, because it can be taken as a disjunction-member of

a conjunction which has only one member. It is a theorem, because p occurs both with and without a curl. Observe that this theorem proves that any proposition can be deduced from an inconsistent set of postulates. For if a set is inconsistent, both p and $\sim p$ are deducible. But in conjunction with our theorem, p and $\sim p$ give q, which stands for any proposition.

§ 3. The Calculus of Pure Logic

The calculus of unanalysed propositions is a part of a system of pure logic which also contains formulas with constituents interpretable as *predicate-variables* (of the first order and type). If ϕ be such a constituent, the postulates of the system, in addition to (*a*), (*b*), (*c*), (*d*), of § 2, are:

$$(e)\ [(x) \cdot \phi(x)] \supset \phi(y);$$
$$(f)\ \phi(y) \supset [(\exists x) \cdot \phi(x)].$$

Besides these new postulates and an obvious extension of the rule of substitution to cover cases in which individual and predicate variables occur, the calculus of pure logic has two additional rules of inference:

(Υ 1) From "$p \supset \phi(x)$" one can derive "$p \supset [(x) \cdot \phi(x)]$";

(Υ 2) From "$\phi(x) \supset p$" one can derive "$[(\exists x) \cdot \phi(x)] \supset p$".

This form of a postulational system of logic can be interpreted as a logic of propositions with no bound-variables except individual-variables. Let us designate it by the initials *P. L.*

The proof of consistency of *P. L.* depends on the possibility of "reducing" its deducible formulas to the theorems of the calculus of unanalysed propositions. If such a "reduction"—to be called "*p*-reduction"—is possible for all deducible formulas of *P. L.*, their contradictions cannot be *p*-reducible or else some of the theorems of the calculus of unanalysed propositions would have to be inconsistent with one another, which as already proved is not true. Hence if all deducible formulas of *P. L.* are *p*-reducible, *P. L.* must be a consistent system.

Now a formula of *P. L.*, applied to a domain of *k* individuals (where "*k*" symbolizes some positive number), is called a "*k*-formula" if it is *p*-reducible. A *p*-reduction of a *k*-formula takes the following steps. First, the individual-variables of the original formula are replaced by values out of the domain of *k* individuals, $a_1, a_2, \ldots, a_k$; second, the prefixes of generality are eliminated by introducing instead conjunctions or disjunctions taken over the domain of *k* individuals; third, each of the propositional functions with arguments is replaced by a different proposition-variable.

To illustrate, let the original deducible formula be:

$$(x) \,.\, [\phi\, x \supset ((\exists\, y) \,.\, \phi\, y)].$$

Expansion by means of conjunction and disjunction gives:

$$[\phi\, a_1 \supset (\phi\, a_1 \vee \ldots \vee \phi\, a_k)] \ldots\ldots [\phi\, a_k \supset (\phi\, a_1 \vee \ldots \vee \phi\, a_k)].$$

After the substitution of proposition-variables for the functions the formula becomes:

$$[p_1 \supset (p_1 \vee \ldots \vee p_k)] \ldots\ldots [p_k \supset (p_1 \vee \ldots\ldots \vee p_k)],$$

which is a theorem of the calculus of unanalysed propositions. This *p*-reduction shows that the original expression is a *k*-formula.

The consistency of *P. L.* can be proved because it is easy to prove that all its theorems are *k*-formulas. The theorems of the calculus of unanalysed propositions are seen to be *k*-formulas by immediate inspection. The postulate (*e*) is *p*-reducible to the theorem that a conjunction of proposition-variables implies one of them; and the postulate (*f*) is *p*-reducible to the theorem that a proposition-variable implies a disjunction of which it is a constituent. Thus both (*e*) and (*f*) are *k*-formulas. As to the rules (Υ_1) and (Υ_2), they are means of deriving *k*-formulas from *k*-formulas because they themselves correspond, respectively, to the theorems (of the propositional calculus) that "$p \supset (q_1 \,.\, q_2 \ldots . q_k)$" is deducible from the premises "$p \supset q_1$", "$p \supset q_2$", . . ., "$p \supset q_k$"; and that "$(q_1 \vee q_2 \vee \ldots \vee q_k) \supset p$" is deducible from the premises "$q_1 \supset p$", . . ., "$q_k \supset p$".

While the problem of consistency is solved by proving that every theorem is a *k*-formula, the reverse, i.e. a proof that a *k*-formula is a theorem, would solve the decision-problem, since one can always test whether a given formula is a *k*-formula by the method of *p*-reduction. In conformity with this consideration the decision-problem has been solved for several kinds of formulas of *P. L.*

First, k-formulas without bound variables are theorems because they can be derived by *substitution* [in accordance with the rules of *P. L.*] from the corresponding theorems of the calculus of unanalysed propositions.

Second, the decision-problem is solved for k-formulas with existential prefixes followed by expressions which contain k free variables. As an example, let the k-formula be:

$$(\exists z) . f(x, y, z).$$

In a domain of two individuals this formula gives:

$$f(x, y, x) \vee f(x, y, y),$$

which, as a 2-formula without bound variables, is a theorem. Let this theorem be a premise in conjunction with two others, which are established as special forms of postulate (f):

$$f(x, y, x) \supset [(\exists z) . f(x, y, z)];$$
$$f(x, y, y) \supset [(\exists z) . f(x, y, z)].$$

From these three premises the original k-formula is deducible and therefore it is a theorem.

Third, k-formulas of the second kind but preceded by non-existential prefixes for all its k variables are also theorems. This is so because the formulas:

$$(\exists y_1) \ldots (\exists y_r) . \phi(a, b, \ldots, k, y_1, \ldots, y_r);$$
$$(x_1) \ldots (x_k) (\exists y_1) \ldots (\exists y_r) . \phi(x_1, \ldots, x_k, y_1, \ldots, y_r),$$

can be proved to have equal deducibility; i.e. when one is a theorem so is the other.

Fourth, k-formulas of the so-called "unary calculus" whose constituent functions have each at most one argument are theorems because they can be reduced to formulas of the third kind. This kind of reduction is a transformation of the original expression into its "prenex" normal form, i.e. a form in which all prefixes precede a matrix which is derived by substitution from the corresponding conjunctive normal form of the calculus of unanalysed propositions.

There is no need to proceed into further details for a general account of the decision-problem. In principle the problem remains unsolved, because there are k-formulas which are not theorems of *P. L.*: these are contradictories of formulas which can be satisfied only in an infinite domain of individuals. Consider the formula:

$$(F)\ (x) . \sim \phi(x, x) . \{[(x, y, z) : \phi(x, y) . \phi(y, z)] \supset \phi(x, z)\} . (x) (\exists y) . \phi(x, y).$$

Its interpretations will show that it cannot be satisfied by any finite number of individuals. For example, if "$\phi(x, y)$" is interpreted as "the integer x is smaller than the integer y", then it is not true, except for an infinite number of integers, that "$(x)(\exists y) . \phi(x, y)$" holds, i.e. that for every integer x there exists a greater integer y. Since the formula (F) cannot be true in any finite domain of individuals, its *contradictory* must be true for all finite domains of individuals and so must be a k-formula. Nevertheless this k-formula is not a theorem, since it fails when k is an infinite number.

§ 4. Gödel's Contribution

When logic is combined with mathematics, either in the logistic manner of the *Principia* (where mathematical notions and statements are resolved into purely logical constituents) or by the addition of mathematical axioms to the postulates of the pure calculus, it can be shown that the combined system contains "undecidable" propositions, is incomplete, and that its consistency is not provable. These are the "negative" results of Gödel's work. At the same time, and in the course of his argument, Gödel has established that the unqualified "vicious-circle" principle is wrong and that propositions within a sufficiently rich language can be about themselves unless they qualify themselves as true, false, or in some other epistemological way.

Gödel's procedure can be divided into three main stages, to be designated as stage I, II, and III. In I a formal system, comprehensive enough to be interpreted as a combined logic and arithmetic, is introduced. This formal object-system will be referred to by the initials *F. S.* In II the system *F. S.* is conventionally represented, we shall say "*arithmetized*", by another system in terms of positive integers. In III special expressions of *F. S.* are constructed with the aid of the arithmetized representations, and these expressions are then interpretable as describing themselves. Some of them are

found to be undecidable formulas. We now shall go over these three divisions at a greater length.

I

The formal object-system *F. S.* can be built up by adding to the pure calculus of logic of § 3 a few primitive elements and postulates which have an arithmetical interpretation. The additional primitive elements (or undefined terms) are:

(1) O (to be interpreted as "the number of zero").

(2) N (to be interpreted as "the successor of" a given number).

(3) $\mathcal{E}$ (to be used as a prefix in expressions of the form "$(\mathcal{E}\, x) . \phi\, x$" which stands for "the smallest integer x such that $\phi\, x$, if there exists an x such that $\phi\, x$; otherwise . . . zero").

For convenience of abbreviation the numbers O, $N(O)$, $N(N(O))$, etc., will be written as z_0, z_1, z_2, etc. Thus z_6 is an abbreviation for "six". The abbreviations in terms of the z's will be called "*transcriptions*". In describing *F. S.* we shall make use of another abbreviation. Instead of writing "the expression obtained from ϕ by substituting the letter a for each occurrence of the free variable x within ϕ" we shall write "Subst $(\phi\, ^{x}_{a})$".

The postulates of *F. S.* (in addition to the postulates (*a*), (*b*), (*c*), (*d*), (*e*), (*f*) of § 3) are:

(*g*) $x = x$.

(*h*) $x = y \,.\, \equiv \,.\, \phi(x) = \phi(y)$.

(*k*) $(x = y) \,.\, (y = z) \,.\, \supset \,.\, x = y$.

(*l*) $\sim (O = N(x))$.

(*m*) $N(x) = N(y) \,.\, \supset \,.\, x = y$.

(*n*) The Principle of Mathematical Induction: From $\phi(o)$ and $\phi(y) \supset \phi N(y)$, $\phi(y)$ can be deduced.

The postulates (*g*), (*h*), and (*k*) determine the properties of the sign "=" to be interpreted as numerical equality: a number is equal to itself; when *x* is equal to *y*, one can replace the other in any context ϕ, and *vice versa*; equality is a transitive relation.

The interpretation of (*l*) is that zero has no predecessor: the system *F. S.* is concerned with positive integers. According to (*m*) no two numbers have the same successor.

In addition to the rules of procedure of the pure calculus the system *F. S.* contains rules of operation with the prefix Ɛ which are entirely analogous, *mutatis mutandis*, to the rules which regulate the use of the existential and non-existential prefixes.

This sums up the formal properties of *F. S.*

II

The primitive elements of *F. S.* are discrete entities, they are countable. And any written expression of *F. S.* must obviously be a finite sequence

of primitive elements. Hence it is possible to assign arbitrarily, as a label, a different positive integer to every different expression (whether it is a primitive element or a combination of such) in *F. S.* This labelling of symbols in *F. S.* is the first step of *arithmetization.*

To arithmetize the *primitive elements* we write under each of them its representing number:

0	N	$=$	$\sim$	$\vee$	$.$	$\supset$	$\equiv$	$\exists$	Ɛ	(	)
1	2	3	4	5	6	7	8	9	10	11	12

Any integer > 13 and $\equiv 0 \pmod 3$, i.e. any integer which being divided by 3 gives no residue, such as 15, 18, etc., will be employed to label *proposition-variables* *p*, *q*, etc.

An integer > 13 and $\equiv 1 \pmod 3$, such as 16, 19, etc., will represent *number-variables*, *x*, *y*, etc.

An integer > 13 and $\equiv 2 \pmod 3$, such as 17, 20, etc., will label a function-variable, ϕ, ψ, etc.

Obviously a *formula*, as a sequence of primitive elements, can be arithmetized by a sequence of the numbers representing these elements. But it is desirable to label each formula by a single number. Let the original arithmetization of a formula be a certain sequence of positive integers:

$$k, k, \ldots, k.$$

It can be correlated with a *single number* defined as the product:

$$2^{k_1} . \, 3^{k_2} . \ldots . \, p_n^{k_n},$$

where p_n is the nth prime number in their order of magnitude.

For example, take the formula in *F. S.*:

$$x = 0$$

As a sequence it is arithmetized into:

$$16, 3, 1.$$

To give it a single number as a label, we must compute the product:

$$2^{16} \cdot 3^{3} \cdot 5^{1} = 65536 \cdot 27 \cdot 5.$$

Proofs are sequences of formulas, of which the last is the conclusion. Accordingly a proof is arithmetized by substituting for each of the constituent-formulas its representative single number. The resulting sequence of numbers can be correlated in its turn with a unique positive integer.

The arithmetization of symbols, formulas, and proofs gives a *one-one correspondence*. Each element or combination of elements in *F. S.* is represented by a different number. Conversely, given the representing integer of a formula, the latter can be "retrieved", because the factorization of a product into its prime factors, which represent the elements of the formula, is unique.

The class of representing positive integers can now be organized into a *system*. The purpose of this organization is to represent within the arithmetized medium *theoretical* (or, according to the postulationalists, metalogical) considerations about the object-system *F. S.*, such as the statement that "a certain

formula of *F. S.* is deducible from another formula of *F. S.*" Thus the arithmetized system consists of *functions* and *relations* of positive integers which can be explained by reference to the theoretical (or metalogical) functions and relations of formulas of *F. S.* As an illustrative selection a few symbols of the arithmetized system are given in the left-hand side column; their elucidation in terms of the theory of the system *F. S.* is given on the right-hand side.

Each of these symbols of arithmetization can be defined directly, i.e. without reference to *F. S.* The direct definitions given in Gödel's original article, although cumbersome technically, have the advantage of showing that all the functions and relations of the arithmetized system are *recursive*.*

* "$\phi(x_1, \ldots, x_n)$ shall be said to be *recursive* with respect to $\psi(x_1, \ldots, x_{n-1})$ and $\chi(x_1, \ldots, x_{n+1})$ if, for all natural numbers

$$\phi(0, x_2, \ldots, x_n) = \psi(x_2, \ldots, x_n);$$
$$\phi(k + 1, x_2, \ldots, x_n) = \chi(k, \phi(k, x_2, \ldots, x_n), x_2, \ldots, x_n)."$$

(K. Gödel, *On Undecidable Propositions*, 1934, Princeton.)

The pair of equations gives a recursive definition of the function. In special cases of recursive definition any variables on the right side of the equations can be omitted in any of its occurrences; in the simplest case the right side of the first equation is a number as in:

$$f(1) = 1;$$
$$f(n + 1) = f(n) \cdot (n + 1),$$

which gives a recursive definition of the function:

$$f(n) = 1 \cdot 2 \cdot \ldots \cdot n$$

A recursive function is computable, i.e. replaceable by a number, in a finite number of steps, because it must be either a function of *1* or of some other number $(n + 1)$. If it is a function of *1*, the first equation of its definition immediately gives its numerical value. If

[*Continued on page* 111

$Neg\,(x)$	The number which arithmetizes $\sim p$, if x arithmetizes p, where p and $\sim p$ are formulas of *F.S.*; otherwise it is zero.
$Sb\left(x^{a}_{N_y}\right)$	The arithmetization of Subst $\left(\phi^{w}_{z_y}\right)$, if x, N_y, and a are, respectively, arithmetizations of ϕ, z_y, w.
$Sb\,(x, y)$ (An abbreviation of the preceding symbol)	
$n\ G\ L\ x$	The nth member of the sequence of positive integers correlated with the product x. Let the sequence be: $k_1, k_2, \ldots, k_k, \ldots, k_n, \ldots, k_r$. Then: $x = 2^{k_1} \ldots p^{k_n}_m \ldots p^{k_r}_q$; and $n\ G\ L\ x = k_n$.
$L\,(x)$	The number of members in the sequence arithmetized by x.
$P\,(x)$	x arithmetizes a proof in *F. S.*
$x\ P\ y$	x arithmetizes a proof of a formula which is arithmetized by y.
Consist.	*F. S.* is a consistent system.

III

At this stage of the argument we shall use "transcriptions" of the expressions of *F. S.* in terms of the z's with subscripts. Let $\phi\,(m, n, \ldots) = k$, where $m, n, \ldots$, and k are positive integers.

Then $\phi\,(m, n, \ldots)$ is transcribed into $g\,(z_m, z_n, \ldots)$ provided $g\,(z_m, z_n, \ldots) = z_k$. A relation

Continued from page 109]

it is $\phi\,(n + 1)$, where $n \neq 0$, it is computable as the function ψ of two arguments which, eventually, are shown to be numbers. For n stands for a number, while $\phi\,(n)$ is either the number a, if $n = 1$, or it is the function ψ of two arguments, n and $\phi\,(n - 1)$. The regression in computation is from $\phi\,(n + 1)$, through $\phi\,(n)$, to $\phi\,(n - 1)$, and so on until the argument of ϕ is reduced to 1, which is bound to happen no matter how large n is.

Let "summation" be the function to be written as "sum $(k, x_2, \ldots, x_n)$."

The recursive definition of "summation" is given, for two numbers, by the equations:

sum $(0, y) = y$;

sum $(N\,(x), y) = N\,(\text{sum}\,(x, y))$, where N stands for "successor".

Let x and y be, respectively, 2 and 1. Then the repeated application of the second equation gives:

$$\begin{aligned} \text{sum}\,(3, 1) &= N\,(\text{sum}\,(2, 1)\,) \\ &= N\,(N\,(\text{sum}\,(1, 1)\,)\,) \\ &= N\,(N\,(N\,(\text{sum}\,(0, 1)\,)\,)\,). \end{aligned}$$

By means of the first equation of the recursive definition the last expression is transformed into "$N\,(N\,(N\,(1)\,)\,)$" which gives 4 as the value of "sum $(3, 1)$".

A *relation* R of positive integers $x_1, \ldots, x_n$, to be written as "$R\,(x_1, \ldots, x_n)$" is recursive if its "associated" function ϕ is recursive. A function ϕ is "associated" with the relation R when the following conditions are satisfied: $\phi\,(x_1, \ldots, x_n) = 0$, if the relation R holds for the same numbers, and $\phi\,(x_1, \ldots, x_n) = 1$, if R does not hold, i.e. if $\sim R$.

$R(m, n, \ldots)$ is transcribed by means of transcribing its "associated" function.*

Now Gödel has proved (and the reader is referred for the proof to Gödel's original paper) that all *recursive* functions and relations of positive integers can be transcribed in terms of z's. But the symbols of the arithmetized system stand for recursive functions and relations of the representing integers. Hence they can also be transcribed, which means that they can be formulated within *F. S.* And since all of them (in their theoretical explanation or import) refer to *F. S.*, in transcription they become expressions of *F. S.* which are about expressions of *F. S.* In special cases some of them are statements about themselves, just as the syntactical statement that "Every English sentence contains a verb" happens to be in English, and therefore is about itself.

Consider the formula $U(w)$ of *F. S.* in the construction of which two other formulas of *F. S.* are employed: $D(u, v)$ which is the transcription of the arithmetized relation $x\,P\,y$; and $S(u, v)$ which is the transcription of the symbol $S\,b\,(x, y)$. The definition of $U(w)$ is given below on the first line of the right-hand side column; its representing number or arithmetization is given on the same line on the left-hand side.

p	$U(w) = (v) \,.\sim D(v, \mathrm{S}(w, w))$. Def.
$\mathrm{S}\,b\,(p, p)$	$U(z_p)$

* For the definition of an "associated" function see the last paragraph of the preceding footnote.

Thus p is the representing number of $U(w)$. If we substitute z_p for w in $U(w)$, the result is symbolized by "Subst $\left(U(w)^{w}_{z_p}\right)$", and the representing number of this symbol is $S\,b\,(p, p)$. The expansion of $U(z_p)$, in accordance with the definition of $U(w)$. gives:

$$(1).\quad U(z_p) = (v) \,.\sim D(v, S(z_p, z_p)).$$

We can now show that $U(z_p)$ is interpretable as a formula which is about itself. First, let the definitions of $U(w)$ be abbreviated as $F(S(w, w))$ and let one of the arbitrary interpretations of $F(z_n)$ be that "the number which is transcribed as z_n, i.e. the number n, arithmetizes a formula which has the property f". Next, we determine the number which is transcribed by $S(z_p, z_p)$; it is "the number of the formula which results from the formula whose number is p when z_p is substituted for its free variable", i.e. it is "$S\,b\,(p, p)$)". This number represents a formula which has, so we learn from our interpretation of $F(S(z_p, z_p))$, the property f. But since this number is the number of $F(S(z_p, z_p))$, i.e. of $U(z_p)$, the latter ascribes to itself the property f.* Thus $U(z_p)$ is an expression which is about itself. This invalidates Russell's "vicious-circle" principle. At the same time it gives an instance of an *undecidable* proposition, i.e. it can be neither proved nor disproved.

Suppose $U(z_p)$ is provable. Then there exists (in

* If we let f stand for "false", then $F(S(z\,p, z\,p))$ becomes a formulation of "This proposition is false", which is intended to apply to itself.

the arithmetized system) a number k such that $k\ P\ S\ b\ (p, p)$. In transcription this gives:

(2) $D\,(z_k,\ S\,(z_p,\ z_p)\,)$.

On the other hand it follows from (1) that, for the value k of v:

(3) $\sim D\,(z_k,\ S\,(z_p,\ z_p)\,)$.

But (2) and (3) contradict one another. Hence if $F.\ S.$ is a consistent system, $U\,(z_p)$ is not provable.

Suppose $U\,(z_p)$ is refutable, i.e. $\sim U\,(z_p)$ is provable. This would mean that $\sim (v)\,.\,(\sim D\,(v,\ S\,(z_p,\ z_p)\,)$, i.e. there exists a number k for which $D\,(z_k,\ S\,(z_p,\ z_p)\,)$ holds. On the other hand, since $U\,(z_p)$ is not provable, $\sim D\,(z_k,\ S\,(z_p,\ z_p)\,)$ should hold for all k. Thus the supposition that $U\,(z_p)$ is refutable also leads to inconsistency.

The fact that $U\,(z_p)$ is undecidable can be used to show that there is no proof of the consistency of the system $F.\ S.$ For if $F.\ S.$ is consistent, $U\,(z_p)$ is not provable. In the arithmetized system this gives:

$$\text{Consist}\,.\supset [(x)\,.\sim (x\ P\ S\ b\ (p, p)\,)]$$

The transcription of this implication is provable. And if the transcribed "consist" were provable $U\,(z_p)$, the transcribed consequent would be provable, and this, we know, is not the case unless $F.\ S.$ is inconsistent.

Thus neither the problems of consistency and completeness, nor the decision-problem, can be solved for logic which is combined with arithmetic.

These so-called negative results have forced the postulationalists to admit that a formal logic cannot be a comprehensive system, that a common language such as English is comprehensive at the price of being inconsistent, and that there is an unending hierarchy of consistent languages arranged in the order of increasing comprehensiveness.

To an intuitionalist Gödel's results are negative in a different sense. They show that postulational systems are always inadequate as expressions of the logic of intuition. Whether this is so because formulations in terms of clear-cut symbols are too stiff to do full justice to involved ramifications and flexible turns of logical thought, is a matter for general speculation. To be more specific, one might argue against mixing up formulas of logic and mathematics. For so long as logic is kept clear of infinite domains, the decision-problem together with the problems of consistency and completeness are solved. And if trouble begins with the infinite, it is bound to come when arithmetic of positive integers, which are infinite in number, is joined with logic. On the other hand, one might look for a deeper source of evil. One might, for example, suspect that the undecidable formula $U\,(z_p)$ is "about itself" in a sense which needs further analysis. This formula has a certain representing number, and when interpreted, it refers to itself as to "the formula which has that number"; but this reference is a definite description, and the question of how the reference by description is possible is far from being explored. Let $U\,(z_p)$ be irreproachable within the abstract framework of a postulational system or even as a mathematical formula, in a larger context of interpretation its "reference to itself" may be a confusion between proposition and propositional function, as indeed one of the interpretations of $U(z_p)$, "This proposition is false", was shown to be.

Chapter IV

CONCEPTUAL REFERENCE

§ 1. Introduction

Objective reference is an element of significance and not of meaning. This is so because of the fact that while the connotative content changes from one proposition to another, all of them are invariably about something. This something is objective in the sense that a proposition which is about it, unless purely verbal, does not refer to a mere word or even to the connotation of the word but to a thing. For example, when I say that I am fond of tennis, I do not mean that I am fond of the word "tennis" or of its definition, I am concerned with the game itself as an actual exercise and enjoyment. It would seem that in speaking or writing about things, one is in contact with extra-linguistic actuality. Hence arises the Paradox of objective reference: "There exists *within* discourse an objective for reference the nature of which is to be something *outside* discourse."

This Paradox is not avoided by treating language as the result of conventions of formation and transformation of sentences. However conventional the basis of a language may be, it must allow, to use Carnap's terminology, for the distinction between real object-sentences and pseudo-object-sentences. Let us illustrate this distinction by Carnap's own examples: he contrasts "Babylon was a big town" as