by the same author

THE PHILOSOPHY OF RELATIVITY

THE PROBLEMS OF LOGIC

by

ANDREW PAUL USHENKO

London

GEORGE ALLEN AND UNWIN LTD

entities which are connotatively distinct from one another. But much as the theory of logic differs from the formalized calculus and much as the difference may remind one of the separation between the metalogic and object-logic, in the intuitional theory there is no real separation because the logical theory is a description of the same thing of which the formulas give an exhibition, viz. of the form of the actual demonstrative thought. For example, the principle that "If one proposition implies another, then the falsehood of the latter implies the falsehood of the former" describes the same thing which is shown by the formula "$p \supset q . \supset . \sim q \supset \sim p$", and both are in their own ways concerned with the form of such concrete statements as "If aggression is provoked by and thus presupposes tolerance of pacifism, then elimination of pacifism would be sufficient to eliminate aggression". And even when in the theory of logic one goes beyond discussions which are descriptive equivalents of symbolic formulas, perhaps beyond statements which could be formalized, the standpoint still remains extensional because however informal the ramifications of a theory may be, they are all ultimately converging upon an explanation of the properties of logical form, which were, to begin with, exhibited for inspection.

Chapter II

THE PARADOXES OF LOGIC

§ 1. Introduction

Traditional logic places no restriction upon premises which are allowed in deduction unless they are downright contradictions. It is confident that no inconsistency can result in deriving a conclusion in accordance with intuitively certain principles from non-contradictory premises. That is where traditional logic is wrong. Logical paradoxes show inconsistency to be an outcome of certain unrestricted formulations, even when these are seemingly tautological or analytic definitions, i.e. explications of the connotation of some given term. Consistency alone, without regard for the conditions of formation of the premises, is insufficient to insure discourse from contradiction. Modern logic is superior to its traditional predecessor primarily because it has realized the necessity for restrictive conditions of formation, a requirement for the consideration of *significance*. Of course, there has always been an instinctive rejection of certain formations of terms as insignificant; for example, in denouncing "Justice is triangular" as an expression which is neither true nor false but meaningless. Yet this sense for discrimination of significance did not find explicit recognition as a principle among traditional logicians; hence they were helpless in facing logical paradoxes. Russell's theory of types, as the first systematic treatment of the paradoxes, was a break with tradition and an introduction to logical restrictions of significance.

A basic distinction of significance is between terms or words which are meaningful in isolation and "incomplete symbols" which are merely contributions to meaning and therefore can be understood only in a context. Proper names are examples of isolatable symbols: they stand for individuals or particular presentations. Descriptions (characteristics and relations), on the other hand, are generally contributions to the formation of a propositional structure, the presence of which they outline, but apart from this contribution to a structure they would seem to have no meaning.

"Attributes and relations, though they may be not susceptible of analysis, differ from substances by the fact that they suggest a structure, and that there can be no significant symbol which symbolizes them in isolation. All propositions in which an attribute or a relation *seems* to be the subject are only significant if they can be brought into a form in which the attribute is attributed or the relation relates. If this were not the case, there would be significant propositions in which an attribute or a relation would occupy a position appropriate to a substance, which would be contrary to the doctrine of types. Thus the proper symbol for 'yellow' (assuming for the sake of illustration that this is an attribute) is not the single word 'yellow', but the propositional function 'x is yellow', where the structure of the symbol shows the position which

the word 'yellow' must have if it is to be significant."*

The emphasis on significance is a point where Russell breaks completely with the traditional logic of terms; the units of his logic are not single entities, but *propositional functions*, complex expressions of the form "x is ϕ", or, more generally "ϕx", where "x" is the *argument* and "ϕ" the *predicate* (descriptive constituent) of the propositional function. One might think that the notational distinction between the argument and the predicate makes a tacit allowance for the entertainment of the predicate as an isolatable meaning. And one can cite other examples, in which there is an intention to be concerned with characteristics as such rather than with propositional structure. To give one, "I prefer 'green' to 'yellow' " is not intended to mean "I prefer the propositional function 'x is green' to the propositional function 'x is yellow' ". To account for this use of predicates Russell introduces the symbol "$\phi\hat{x}$" (read "phi-ex-cap") which is not a propositional form and therefore, unlike the symbol "ϕx", cannot be transformed into a proposition.† Nevertheless the form "ϕx" remains fundamental in the *Principia* in the sense that it determines the significance of "$\phi \hat{x}$" as well as of any other logical construction, including propositions.

* *Contemporary British Philosophy*, First Series, p. 375. (George Allen & Unwin Ltd.)

† For an illuminating discussion of the purely logical basis of the distinction between these two kinds of function cf. ch. 1 of the mimeographed *Mathematical Logic*, by A. Church.

For a treatment of logical paradoxes the dependence of the significance of propositions on the significance of functions is of primary importance. This dependence can best be explained by introducing the two ways of transformation of a propositional function into a proposition, *evaluation* and *generalization*.

To illustrate these transformations we shall use a propositional function with a fixed meaning, "*x* is destructible", for which the symbol "*f x*" will be an *abbreviation*. The process of evaluation is a derivation of a proposition by means of assigning some *value*, i.e. some fixed meaning, to the argument of the function. For example, if we assign to the argument "*x*" the meaning "Rome", the evaluation gives the proposition *Rome is destructible*, to be abbreviated as "*f a*".* The same propositional function can also serve as a basis for generalization over the argument, one kind of which gives an existential and another a non-existential proposition, viz. $(\exists x) . f x$,

* The notation used in this book does not differ essentially from the symbolism of the *Principia Mathematica*. But in addition to the convention of using the letters of the beginning of the alphabet as names for individuals and the letters "*x*", "*y*", "*z*" for individual-variables, we shall confine the letters "*f*", "*g*", "*h*" to abbreviations of predicates having a fixed meaning. Such an abbreviation of a proposition as "*fa*" is to be distinguished from the function "*fx*" as well as from the function "ϕa", which has a predicate-variable. Without prejudging the question whether a predicate is always a propositional function, we shall use a single letter to designate a predicate as abstracted from its argument. Thus where Russell writes "$\phi \hat{x}$", we shall write instead "ϕ". It is of some interest that in exceptional cases the *Principia* itself adopts this simplified notation. Cf. *Principia Mathematica*, second edition, p. 49.

i.e. *Something is destructible*, and $(x) . f x$, i.e. *Everything is destructible*.

A question of consequence for the theory of logical significance is whether or not transformation by generalization can be reduced to transformation by evaluation. Some logicians have tried to identify an existential proposition with a disjunction and a non-existential proposition with a conjunction of singular propositions derivable by evaluation from the same function. According to them "Everything is destructible" means: *Rome is destructible and Mars is destructible and, etc.* (for all values of "*x*"). But in order to identify a general proposition with a collection of singular propositions not only must the latter be actually enumerated but the very possibility of incompleteness of this enumeration should be excluded. This condition is satisfied when either context or perceptual evidence specify the *number* of values which the argument of the propositional function (in the basis of generalization) can take. Let the context be a discussion of probability:

"We throw a single die of normal unbiassed construction under normal conditions. . . . We know (*a*) that the uppermost faces must be either *one* or *two* or *three* or *four* or *five* or *six*, we know (*b*) that it will not be more than one of these."*

This context authorizes us to say that "All faces of the die have the same probability of falling uppermost, viz. *1/6*", which means, however, nothing more than the conjunction of singular propositions "The face *one* has the probability (of falling upper-

* C. A. Mace, *The Principles of Logic*.

most) 1/6, and the face *two* has the probability 1/6, and the face *three* has the probability 1/6, and the face *four* has the probability 1/6, and the face *five* has the probability 1/6, and the face *six* has the probability 1/6". Instead of a context we sometimes make use of direct perception, as when, entering a room and observing three seats but none of them empty, we say "All seats are occupied", meaning the same as "The first seat is occupied and the second seat is occupied and the third seat is occupied". But, of course, the two statements would not mean the same thing to a person outside the room, to him the conjunction must be conveyed together with the clause of exclusion "And there are no other seats in the room". Without the aid of either perception or context, a general proposition can be reduced to a collection of singular propositions only when the latter are in conjunction with the clause of exclusion "And there are no others" which insures completeness of the enumeration of singular propositions. But the clause of exclusion is itself a general proposition, and therefore our conclusion is that a general proposition as such (i.e. without the aid of additional information from perception or context) is not translatable into singular propositions. This conclusion is strengthened in case the number of singular propositions is too great to be enumerated, or when there are no corresponding singular propositions and the general statement is true vacuously, as is the statement that "All ghosts are transparent".

But while general propositions are irreducible and "generality is seen to be an ultimate mode of significance, in extensional logic, where equivalence of propositions is determined exclusively by the identity of their truth-value, the treatment of a general proposition as a truth-function of singular propositions has at least the justification of a fiction useful for computation." A general non-existential proposition would be false when and only when at least one of the singular constituents of the corresponding conjunction is false, and true otherwise. This condition holds even if the number of the constituents should be infinite. And if there are no constituents at all, the condition cannot be violated. *Mutatis mutandis* an analogous condition can be formulated for the reduction of existential propositions.

§ 2. Russell's Theory of Types

Russell's theory of types is a systematic definition of logical significance of propositional functions by means of restrictions put upon the range of values which can be assigned to their arguments.* The *type* of a logical entity is the mode of significance which it takes in a context. Within a propositional function, the simplest kind of context, the type of the function-variable is determined with regard to the possible values of its argument. Thus the signifi-

* The following exposition is intended to give the essentials of the doctrine of types without going into such details as the distinctions to be found in the two editions of the *Principia*. I believe these distinctions have now merely an historical interest, for the improvements of the second edition were steps in the direction which has been fully explored since by F. Ramsey. His findings were endorsed later by Russell himself.

cance of "human" in the propositional function "x is human" is its capacity to characterize individuals, where anything that can be tagged by a proper name is an individual. With any other object of characterization the propositional function would be transformed into a meaningless expression. For example, "2 plus 2 is human" is neither true nor false, but insignificant. As a general rule a function cannot significantly take itself as a value. This excludes any such expression as "human is human" or, more generally the form "$\phi\ \phi$". Given two functions which take the same values of the same kind and therefore are of the same type, neither can be used for reciprocal evaluation. "Proust is famous" and "Proust is human" are both significant; therefore it is meaningless to say either "famous is human" or "human is famous". We can establish now the distinction between the lowest type, the designations of individuals, and the next higher type, the characteristics of individuals.*

But, of course, characteristics of individuals can themselves be described by certain other characteristics. In general, besides functions of individuals there are functions of functions. "Anything human is mortal" is a proposition derived by evaluation from such a function, namely from the expression "Anything ϕ is mortal", which can be written in the abbreviated form "$F\phi$". The function "F" is said

* We do not say "the designations of characteristics of individuals", because in the realist idiom of the *Principia* entities themselves (with the possible exception of individuals), and not merely their symbolic designations, are classified into types of significance.

to be of the next higher type than the function-variable "ϕ" and two types higher than the individual-variable "x". The rule for type-division is very simple: a function is one type higher than the type of its arguments. A systematic application of this rule gives the hierarchy of individuals, functions of individuals, functions of functions of individuals and so on. We can call a function of individuals a function of type 1; a function of functions of individuals a function of type 2; and so on for the succession of higher types.*

In the broad sense of "type", the word is used not only for the divisions of the above hierarchy but also for the distinctions of *order*. We shall always take "type" in its narrow sense, as contrasted with "order". The hierarchy of orders is built up on the basis of a propositional function of any type by means

* This hierarchy is sometimes called absolute because it begins with the ontological distinction between individuals and characteristics. But Russell himself has suggested that for all purposes of logic a relative hierarchy, where the type of "individuals" is merely a designation for the lowest type relatively to a context, even if the entities of the lowest type are functions, is sufficient. It is not explained, however, how to determine the rank of types in a context. Suppose we have in a given context three entities of the types "d", "e", "f", respectively. We might decide to take "d" as the type of "individuals", "e" as the type 1, and "f" as the type 2 in a relative hierarchy. But it would seem that unless our decision is purely arbitrary, it must be based on some order established outside the context, in our example, on the order of letters in the alphabet, where "e" comes after "d" and "f" comes after "e". And since the order of the alphabet is not constitutive of the type-significance it must be taken as a symbolic scheme of the absolute hierarchy of types.

of generalization, which means that even when functions have arguments of the same type, they may be functions of different orders. Let us take a function which is a *matrix*, i.e. which is free from generalization, "x is greater than y", or, to abbreviate,

$$g(x, y).$$

By generalizing over either of the variables we derive functions other than matrices:

$$(x) . g(x, y); (y) . g(x, y); (\exists x) . g(x, y); (\exists y) . g(x, y).$$

By generalizing the remaining variable in any of these four expressions we obtain propositions. Thus from the first expression we derive:

$$(x)(y) . g(x, y); \; (x)(\exists y) . g(x, y).$$

The original matrix as well as all the derivative expressions and propositions belong to the first order. A logical entity of the first order contains only individual-variables, free or bound. From this definition it follows that a truth-function, of which the constituents are expressions of the first order, is itself a first-order expression. To have a second-order function we must take a matrix in which both the predicate and the argument are variables; we separate the predicate from the argument by an exclamation sign to indicate that the symbol is a matrix:

$\phi \,!\, x$ (read "phi-shriek-ex").

Let this be a function of two variables, "ϕ" and

"x", where the values of "ϕ" are first-order functions, such as "destructible", "greater than", etc. Then the matrix "$\phi \,!\, x$", as well as any expression or proposition derived from it by generalization, is a logical entity of the second order. To understand the significance of the distinction between characteristics of the first and the second order, let us examine the expression:

x has all the characteristics of an artist.

"Having all the characteristics of an artist" is itself a characteristic of an artist, but in a new sense which was not meant in the original meaning of the word "characteristic"; in the new sense it is a function of the second order, originally it was taken as a function of the first order. We have here an illustration of the difference between the forms "$(\phi) . \phi x$" and "$f x$". Both are functions of "x", and the first form can be written as "$F(\phi, x)$", but a restriction of order-significance allows for "f" while it excludes "F" as a possible value of "ϕ" in the matrix "$\phi \,!\, x$". As we have here a distinction between first-order and second-order functions, so we can proceed further with the recognition of functions of still higher orders. The rule that a function-variable can take as values only functions of a lower order than itself determines the clear-cut separation of one order from another.

The hierarchy of types (in the narrow sense) together with the hierarchies of order within each type form the so-called *branched* division of types (in the broad sense). The subsequent developments of the theory of logical significance have, in the main, consisted in efforts to abolish the distinctions of order and to retain only the *simple hierarchy* in which the type of a function is determined by the type of its argument.*

§ 3. The Paradoxes

The paradoxes can be divided into two groups: the logical paradoxes which are resolved by the divisions of type and the epistemological paradoxes which can be dealt with by means of the distinctions of order.

The most famous example of the first group is Russell's paradox of the class of classes which are not members of themselves. In order to avoid the misconception that this paradox is merely a disclosure of an inherent inconsistency in the notion of a class, it is advisable to formulate it in the language of propositional functions.

Functions either can or cannot be predicated of themselves. "Conceivable" is itself a conceivable characteristic. But the property "feline", unlike a

* Even among the recent expositors of the doctrine of types the basic importance of the distinction between types and orders has not always been understood. Thus the account in *Symbolic Logic* by C. I. Lewis and C. H. Langford is virtually worthless because of the insufficient statement on p. 454: "Thus functions of different orders are necessarily of different types, whereas functions of the same order may or may not be of the same type. However, not much emphasis is placed upon this difference of type within the same order, and for all practical purposes functions of the same order can be regarded as being of the same type."

cat, is not itself feline. A function or characteristic which is not predicable of itself may be called impredicable and symbolized by "I". The formularized definition of I is:

$$I(\phi) = \sim \phi(\phi) \;.\; \text{Def.}$$

This definition allows for the equivalence:

$$(\phi) \,.\, I(\phi) \equiv \sim \phi(\phi).$$

Since the equivalence is true for every ϕ, it must be true when ϕ is given the value I:

$$I(I) \equiv \sim I(I).$$

But this result is a contradiction.

Resolution: The simple hierarchy of types rules out any expression in which a function is an argument of itself. Hence the formularized definition of I, in the definiens of which ϕ is the argument of ϕ, must be rejected as meaningless.

The epistemological paradoxes were already known to the ancient Greeks in the form of Epimenides' predicament.

Epimenides, who was a Cretan himself, is supposed to have said that "All Cretans are liars". If we interpret his statement as "All Cretan assertions are false" and *assume*, for the sake of argument, that all other Cretan statements actually were false, then Epimenides' own statement leads to a contradiction. For if it were true, according to its own meaning it should be false along with other Cretan assertions. And if it were false, then there should be some true

Cretan statement, in fact it should be Epimenides' own statement since by assumption the other Cretan statements could not be true. And so "the music goes round and round". Let us formularize the argument. We shall use "*s*" as a statement-variable, "*e*" as a name for Epimenides' assertion, "*f*" as an abbreviation for the characteristic "asserted by Cretans".

(1) $\sim(\exists s).[fs.s.\sim(s=e)]$.	By assumption
(2) $e.\equiv[(s).(fs\supset\sim s)]$.	By definition of *e*
(3) $e\supset(fe\supset\sim e)$.	By (2)
(4) $[(fe\supset\sim e).fe]\supset\sim e$.	By the principle that "$[(p\supset q).p]\supset q$"
(5) fe.	By information
(6) $e\supset\sim e$.	By (3), (4), & (5)
(7) $\sim e\supset[(\exists s).(fs.s)]$	By (2)
(8) $\sim e\supset e$.	By (1) and (7)
(9) $e\equiv\sim e$.	By (6) and (8)

The division of propositions into orders gives an easy solution of the difficulty. Epimenides' assertion, if intended to apply to itself, is meaningless. It can be interpreted as a meaningful statement only if it is one order higher than the Cretan propositions it refers to. But if so, the transition from (2) to (3), whereby *s* takes *e* as a value, is illegitimate. And without this transition there is no contradiction.

Even a more striking epistemological paradox, usually taken to be merely a simplified version of "Epimenides", is the assertion "I am lying" or

"This proposition is false" (when it is intended to apply to itself). Grelling's thorough analysis of it, for which he himself gives credit to Lukasiewicz, is as follows.*

"Let '*q*' be an abbreviation of the phrase 'the proposition on the 8th line of this page'. Then let us write:

(1) *q is a false proposition.*

By counting the lines we verify:

(2) *q is identical with the proposition* (1).

(3) *q is a false proposition* is equivalent to non-*q*.†

The first member of this equivalence (printed in italics) is our proposition (1). Thus we have:

(4) The proposition (1) is equivalent to non-*q*.

But in virtue of (2) *the proposition* (1) can be replaced by *q*. Thus results the contradiction:

(5) *q* is equivalent to non-*q*."

In order to solve this paradox, we observe that if we give to *q* in (1) its value in accordance with (2), we can rewrite (1) as: (1) The proposition (1) is a false proposition.

But if this were a proposition it would mean a violation of the order-prohibition of general propositions (and a singular proposition involving a definite description is a species of a general pro-

* Cf. K. Grelling, "The Logical Paradoxes", *Mind*, 1936.

† This equivalence is assumed to be generally accepted in logic.

position) which apply to themselves. Hence expression (1) is meaningless and not a proposition.

The foregoing exposition conforms with Russell's original view that the same "vicious-circle" principle, which outlaws the application of an expression to itself, resolves both the logical and the epistemological paradoxes:

"Whatever involves *all* of a collection must not be one of the collection; or, conversely: If, provided a certain collection had a total, it would have members only definable in terms of that total, then the said collection has no total."*

But Russell's own disciple, F. Ramsey, has shown that the difference between the logical and the epistemological paradoxes is so essential that they require different treatment. He also raised some doubts about the "vicious-circle" principles, which have been fully justified by recent developments in the formalist logic: there are legitimate expressions which are about themselves.

§ 4. Criticism of the Order-Hierarchy

The main fault with the hierarchy of orders is that it rules out, together with the epistemological paradoxes, certain important propositions which otherwise would seem to be perfectly correct.

A traditional philosophical method of refuting a thesis is by showing that it cannot stand its own test. When the sceptic tells us that nothing can be known,

* Russell and Whitehead, *Principia Mathematica*, second edition, p. 39.

the philosopher answers that if this were so, the sceptic could not know it to be so. Such an answer is effective and philosophers would hate to give it up merely because according to the doctrine of orders no thesis can be self-refuting since no proposition can be about itself.

Yet in the movement against the order-hierarchy it was not philosophers but mathematicians who took the lead. They did it in defence of an important kind of mathematical procedure, which is exemplified in the definition of the least upper bound of a series of real numbers.

The least upper bound of a set of numbers in the ascending order of magnitude is the least number which is not less than any number in the set. If among the numbers of the set there is one which is the greatest, it is called the *maximum*, and, in this case, it is the least upper bound. For example, in the series "1, 2, 3, 4, 5" the upper bound and the maximum is 5. On the other hand, when there is no greatest number in the set, the least upper bound is a certain number outside the set. In the dense series of all rational fractions less than 1, the upper bound is its limit 1, which is the least number of a series of rationals from 1 upward. Now according to Dedekind's postulate, if we divide any ascending series of numbers into two jointly exhaustive and mutually exclusive series, the lower section and the upper section, then there is a dividing number which is the upper bound of the lower section. But it would seem that the series of rationals does not satisfy Dedekind's postulate. For suppose we divide this series into the lower section of all fractions whose square is less than 2 and the upper section of all fractions whose square is greater than 2. Since it can be proved that there exists no fraction the square of which is equal to 2, there is no number which divides the sections unless, to save the postulate, we construct a pure fiction, the existence of the irrational $\sqrt{2}$ as the dividing number. However, there is a better procedure than indulgence in fictions; the series of rationals can be transformed into a new series, which satisfies Dedekind's requirement, the continuous series of positive *real numbers*. The terms of this series are not the rationals themselves but *classes* of positive rationals, and the ascending series of these classes is ordered by the relation of whole and part. To build up this derivative series we define its terms as classes of positive rationals which have no maximum in the original series of rationals. This means that to every rational number there corresponds a term of the derivative. Corresponding to $\frac{1}{2}$ there is "the class of fractions smaller than $\frac{1}{2}$", corresponding to the rational 1 "the class of proper fractions"; and so on. But the derivative series of real numbers also contains terms to which there is no corresponding fraction; one of these terms is "the class of all rationals smaller than $\sqrt{2}$", and it is identified with $\sqrt{2}$. Clearly $\sqrt{2}$ is the least upper bound of all those classes of rationals that correspond to rationals whose square is less than 2. In general an upper bound of a series of real numbers is defined as the real number which is the logical

sum of the real numbers smaller than it. In other words, the members of the upper bound are all those rationals which are members of any class-term of the series concerned. In the language of propositional functions or characteristics we can say that the membership of the upper bound is defined by the characteristic of "having any of the characteristics which determine the membership of those classes which are the terms of the series". But according to the divisions of order the characteristic of having any of the characteristics (of a kind) is of a higher order than the latter. Thus if the latter define real numbers, the former, being higher in order, cannot be a real number. And there the theory of real numbers breaks down.

In order to counteract this adverse effect of the hierarchy of orders Russell offered his *Axiom of Reducibility*:

$$(\exists \psi) . (\phi x \equiv \psi ! x), \text{ i.e.}$$

Given a function of any order ϕ, there exists an equivalent function of the first order ψ, where "equivalent" means that with the same values of the arguments both functions are transformed into propositions which have the same truth-value. With the aid of the Axiom of Reducibility the characteristic of having any characteristic which determines the membership of the class-terms of a series of real numbers can be replaced in the definition of the upper bound by an equivalent characteristic one order lower, and when the upper bound of a series of real numbers is thus redefined it itself can be taken as a real number. This saves the mathematical theory, but raises the question of justification of the Axiom of Reducibility.

There is no doubt that in ordinary language we have no difficulty in reducing the order of characteristics. The characteristic of having any characteristic of an artist is, perhaps, equivalent to the characteristic of a lower order "temperamental" or "observant" or, if this is not so, to a disjunction of first-order characteristics that artists have ever had. A further evidence in favour of the Axiom, at least for those logicians who believe in the existence of classes, is that their belief entails the Axiom, as Russell himself has pointed out. If there are classes, then a function of any order "ϕx" determines a class "C" of those values of the argument of the function which transform the latter into a true proposition. Then we can say "x is a member of C". But this expression is equivalent to "ϕx" and is itself a first-order function.

Sometimes a fear is voiced that the Axiom of Reducibility is in effect a removal of all barriers of type and order, a removal which is at the same time a restoration of the contradictions. This is not so. The Axiom of Reducibility does not affect the simple hierarchy of types, because it is concerned with the reduction of order exclusively. Nor does it reinstate the epistemological paradoxes; for the cause of these is not a mere confusion in taking functions of different orders to be equivalent with respect to their truth-value, but rather the practice of substituting one function for another with no contextual

restriction and in disregard of their difference in connotation. As Ramsey puts it:

". . . this second set of contradictions are not purely mathematical, but all involve the ideas of thought or meaning, in connection with which equivalent functions (in the sense of equivalent explained above) are not interchangeable; for instance, one can be meant by a certain word or symbol, but not the other, and one can be definable, and not the other."*

There remains a general objection, and Ramsey makes much of it, that the Axiom of Reducibility is not an analytic truth or a tautology. This objection may have weight with the members of the logistic school, who want only indubitable principles as a basis for logic, but it should be irrelevant so far as the postulationalists are concerned since they have no regard for axiomatic self-evidence and can accept without scruples the Axiom of Reducibility as one postulate among others.

It is undeniable, however, that the whole construction of the order-hierarchy together with its annex, the Axiom of Reducibility, is extremely cumbersome. Even without the Axiom, one must put up with the awkward condition of an endless duplication, for every order of the hierarchy, of the set of logical principles. For instance let the principle be:

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

According to the doctrine of orders this principle

* F. Ramsey, *The Foundations of Mathematics*, p. 28, Kegan Paul, 1931.

cannot apply to itself. Yet it must be either true or false. Hence we must either subsume it under another principle of the excluded middle, of a higher order, or else, which is much simpler, try to do away with the orders altogether.

§ 5. Rejection of Order-Divisions

The abolishment of the hierarchy of orders is the result of Ramsey's work. He knew that order-ascension depends on generalization, he also knew that in extensional logic general propositions have the same significance as truth-functions of singular propositions; the conclusion that there is no logical significance in the divisions of order was inevitable.

Ramsey's critics, who believe that he overlooked the difference in connotation between a general proposition and the corresponding combination of singular propositions, are entirely wrong. Not only he did not overlook this difference, but, on the contrary, he used it as a guiding principle in his solution of the epistemological paradoxes, which is as follows.

Whatever the relation between a general proposition and the corresponding truth-function, they differ in *formulation*. By means of such linguistic differences one can express different *ways of meaning* things, even when the things referred to are the same. Let two distinct formulas F_1 and F_2 express two different relations of meaning M_1 and M_2, the object of reference being in both cases O, and

"$M(F, O)$" be an abbreviation of "The formula F means by the relation of meaning M the object meant O". Then, it is clear that while "$M_1(F_1, O)$" and "$M_2(F_2, O)$" are true, "$M_1(F_2, O)$" and "$M_2(F_1, O)$" are false. These symbolic illustrations show that a misuse of the relation of meaning generates *falsehood* and not lack of significance. Ramsey concludes that the epistemological paradoxes are instances of such a misuse of meaning and therefore result in falsehood. For example, Epimenides the Cretan must have meant, in a sense which can be called M_2, that "All propositions meant (in a sense to be called M_1) by Cretans are false". There is no paradox unless one confuses M_2 with M_1. But this confusion does not transgress logical significance, it is simply a false description of how Epimenides meant what he said.*

Of course, M_2 may be said to be a higher order of meaning than M_1, and since this would lead to a hierarchy of orders of meaning, the corresponding propositions must also be arranged in a hierarchy of orders. In fact, this gives the same hierarchy as Russell's, with the important difference that it represents variation in formulation and not distinctions in logical form.

"My solutions of these contradictions," says Ramsey, "are obviously very similar to those of Whitehead and Russell, the difference between them lying merely in our different conceptions of the

* This is merely a brief summary of the involved argument to be found in the op. cit., pp. 42–49.

order of propositions and functions. For me, propositions in themselves have no orders; they are just different truth-functions of atomic propositions—a definite totality, depending only on what atomic propositions there are. Orders and illegitimate totalities only come in with the symbols we use to symbolize the facts in various complicated ways."* Ramsey's rejection of order-divisions was accepted almost universally and logicians have concentrated in subsequent writing upon the simple hierarchy of types. The work of the postulationalists has been, in this respect, especially impressive. But I think that the intuitionalist has the right to point out that Ramsey's argument is conclusive only within the logic of propositions. When the elements of logic are merely sentences or strings of marks, the distinctions between the different ways of meaning cannot be made. Of course, this distinction is unnecessary for a postulational system taken in abstraction from its interpretations. The question, however, must be raised whether metalogic or semantics which refer to the object-logic should take account of the various ways of reference. And if it should, one might expect that rejection of propositions as distinct from mere sentences must force the postulationalists to recognize the various ways of reference as so many different *logical* formulations. This would be, in effect, giving up Ramsey's conclusion that the divisions of order have no logical significance.

* Op. cit., p. 48 f.

§ 6. The Postulational Treatment of Types

There is a sharp distinction between the principles of consistency and significance in logistic; even the grounds of their assertion are not the same; the rules of significance are not intuitively certain, they have the air of *ad hoc* devices to escape from the logical paradoxes.

The postulational logic has the advantage of a uniform treatment; since it has no regard for intuitive self-evidence, it can formulate the restrictions of significance in the same way in which it gives the other initial conditions of procedure as so many postulates. Thus it might profit from Russell's simple hierarchy of types by postulating it and then annexing to the other postulates of its system. Such a simple procedure would be advisable, if it were not for the fact that, even when purged from the complications of order, the theory of types is far from being satisfactory. The defects of the simple theory of types are ably summed up in a statement by W. V. Quine:

"But the theory of types has unnatural and inconvenient consequences. Because the theory allows a class to have members only of uniform type, the universal class V gives way to an infinite series of quasi-universal classes, one for each type. The negation $-x$ ceases to comprise all non-members of x, and comes to comprise only those non-members of x which are next lower in type than x. Even the null-class gives way to an infinite series of null classes. The Boolean class algebra no longer applies to classes in general, but is reproduced rather within each type. The same is true of the calculus of relations. Even arithmetic, when introduced by definitions on the basis of logic, proves to be subject to the same reduplication. Thus the numbers cease to be unique; a new *0* appears for each type, likewise a new *1*, and so on, just as in the case of V and Λ. Not only are all these cleavages and reduplications intuitively repugnant, but they call continually for more or less elaborate technical manœuvres by way of restoring severed connections."*

This criticism raises a two-fold problem. On the one hand, since the hierarchy of types leads to "unnatural and inconvenient consequences", it must be dispensed with. On the other hand, formations of symbols which lead to paradoxes must be ruled out without the aid of type-divisions. The postulational logic has all the technical equipment for the solution of this problem. It rejects the restrictions of type by allowing a function to be its own argument. With this allowance there are at least three methods of avoiding the paradoxes, associated in American literature with the names of Alonzo Church, W. V. Quine, and H. Curry, respectively.†

* "New Foundations for Mathematical Logic," p. 78 f. *Am. Math. Monthly*, February 1937.

† In the following sketch no attempt is made to be faithful to the rigour and thoroughness of the original presentations, for which the reader is referred to *Mathematical Logic*, Lectures by Alonzo Church, Princeton University, 1935–36, pp. 16 ff.; *New Foundations for Mathematics*, by W. V. Quine, pp. 77 ff.; and "First Properties of Functionality in Combinatory Logic", by H. B. Curry, *The Tohoku Mathematical Journal*, February 1936.

Church offers a definite criterion for dividing expressions into two groups. He defines one of them as the group of meaningless expressions and shows that it contains the logical paradoxes.

In his system symbolic expressions bracketed after the pattern

(1) $$[---](\ldots)$$

indicate that it is permissible to substitute for the free variable of the function within the square brackets the whole expression in round brackets. The rules of substitution are formalized and their application may or may not terminate in a derivative formula without free variables. If it does, the original formula is said to be meaningful, otherwise it is meaningless. There are no postulates which would rule out as insignificant such a formula as "$\sim \phi(\phi)$", where ϕ is an argument to itself. But take

(2) $$[\sim \phi(\phi)](\sim \phi(\phi)).$$

This expression is bracketed after the pattern (1) and therefore in agreement with the rules of substitutions can be transformed into

(3) $$\sim [\sim \phi(\phi)](\sim \phi(\phi)).$$

But (3) shows the same pattern of bracketing as in (1); and the corresponding substitution leads back to (2). This circle is an expression of Russell's paradox, but (2) is now rejected by definition of a meaningless formula, and not because it involves a function which is applied to itself.

In Quine's procedure, his observation that only one rule of the system can be responsible for the occurrence of Russell's paradox makes him postulate a restriction upon substitution for the free variables involved in this rule, while all other rules remain free from restrictions.

Quine shows that Russell's paradox in the form

$$(\exists x) . [(x \epsilon x) \equiv \sim (x \epsilon x)]$$

is a special case of the theorem

$$(\exists x) (Y) . [(Y \epsilon x) \equiv \sim (Y \epsilon Y)],$$

which can be derived in his system by the rule *R* 3: If "*x*" does not occur in ϕ, $(x) (Y) : (Y \epsilon x) \equiv \phi$ is a theorem. The introduction of a restriction is said to *stratify* ϕ, but only within the context of *R* 3. In accordance with stratification, whenever ϕ is a complex involving the relationship of membership ϵ, the variable on the left of ϵ must be one type lower (using the terminology of the theory of types) than the type of the variable on the right side. To quote Quine:

"I will now suggest a method of avoiding the contradictions without accepting the theory of types or the disagreeable consequence which it entails. Whereas the theory of types avoids the contradictions by excluding unstratified formulas from the language altogether, we might gain the same end by continuing to countenance unstratified formulas. Under this method we abandon the hier-

archy of types, and think of the variables as unrestricted in range. But the notion of stratified formula . . . survives at one point: we replace *R* 3 by the weaker rule:

R 3′. If ϕ is stratified and does not contain '*x*',

$(\exists x) (Y) (Y \epsilon x) \equiv \phi)$ is a theorem."*

Finally, Curry offers a formal method for determining the logical categories of various expressions. A function applied to itself—which can be written in Curry's simplified notation of rows of entities as *ff*—gives rise to Russell's paradox only if it comes under the category of a proposition, but nothing in the rules of his system indicate that it should be so interpreted.

Curry shows that the form of Russell's paradox

$$ff = \sim (ff)$$

can be proved in his system by a joint use of the operators *W* and *B*. A row of two entities prefixed by *W* give the original row with the last entity duplicated. Thus

(1) $$W (B N) f = B N f f.$$

The result of prefixing a row of three entities by *B*, is the bracketing of the last two entities of the row as a single expression.

(2) $$B N f f = N (f f).$$

Since *f* and *N* may be anything, let *f* be defined as

* Op. cit., p. 79.

$W\,B\,N$ and N as the sign of negation "$\sim$". These definitions transform (1) and (2) into

$$ff = B \sim ff;$$
$$B \sim ff = \sim (ff),$$

from which it follows that

$$ff = \sim (ff).$$

Curry then shows that in his system ff may not be a proposition and therefore the last equation is not a contradiction.

All three methods seem to be formally correct. This however, cannot be ascertained before it is proved that they do not lead to inconsistency within and in conjunction with the postulational systems to which they belong. Thus in the postulational logic the problem of significance becomes an aspect of the problem of consistency. Negatively this is shown in all three methods sketched above since they have no use for the distinctions of type-significance and allow a function to take itself as a value of its argument. Positively the tendency to disregard significance (as distinct from consistency) by making it a matter of interpretation is most pronounced in Curry's procedure. But while the tendency to reduce logical difficulties to questions of consistency can be praised as leading to basic simplification, so far as the paradoxes are concerned it means a postponement of their solution. First, because the solution must wait for the proof of the consistency of the system; secondly, because the proofs of consistency

are metalogical discussions in which reference to certain formulas of the object-system revives the epistemological paradoxes. And since the postulationalists cannot resort to Ramsey's mere linguistic distinctions of order as opposed to distinctions of significance, they end by making a much more drastic distinction of languages which are irreducible to one another. Thus, according to Gödel, any language B, however rich, cannot contain such expressions as "false statement in B"; these expressions must belong to a meta-language. But surely one does talk in English about "false English statements". The postulationalists declare that this proves that English is self-contradictory! To my mind to say that English is self-contradictory is to indulge in a paradox far more intolerable than "Epimenides" or any other epistemological puzzle; it certainly is not giving a solution of the puzzles.

§ 7. An Individual Examination of the Paradoxes

The strength of the postulational logic—the explicit statement and formularization of all the postulates and rules in use in the system—is more likely than anything else to be in the long run a weakness in its dealing with the paradoxes. For, as a fixed set the postulates of a system may be too rigid, even if they have disposed of all difficulties up to date, to meet the emergency of some new puzzle or even of some original version of an old one. After all Ramsey made a step forward when he understood that there is no single principle, such as

Russell's "vicious-circle" principle, which would do away with all logical difficulties. But perhaps his own division of paradoxes into two groups was still too summary as a way for complete success. And if so, it would seem to be profitable to relinquish any set-up mechanism, but approach the paradoxes individually, as they come out, and without the prejudice that the method of dealing with one of them would be effective with another. Accordingly and in contrast with the postulationists who operate with a fixed number, however great, of rules, I shall examine each paradox with the possibility in mind that it might require a unique treatment or even suggest the formulation of a principle which has never been thought of before. Of course, this attitude marks a belief in the flexibility and resourcefulness of logical intuition.

I shall begin with my own version of the "Epimenides" which is so formulated as not to be amenable to the kinds of solution which are given by either Russell or Ramsey.

All propositions written within the rectangle of Fig. 1 are false.

FIG. 1

Let the expression within the rectangle of Fig. 1 be called a and let f denote the phrase "written within the rectangle of Fig. 1." Then:

(1) $f\,a,$

(2) $\sim (\exists p) . [f\,p \,.\, \sim (p = a)].$

(3) $a = (p) . (f\,p \supset \,\sim p).$

Suppose a is itself a proposition. Then if a is true:

$$f\,a \supset \,\sim a, \quad \text{by (3)}$$
$$f\,a$$

(4) $\sim a$

But if $\sim a$, then (3) gives:

$$(\exists p) . (f\,p \,.\, p).$$

This result is compatible with (2) only if:

(5) $a.$

The vicious circle—from (4) to (5) and back—cannot be avoided if, following Russell, we declare that a, intended to apply to itself, is not a proposition but a meaningless expression. If a is not a proposition, then:

$$\sim (\exists p) . f\,p, \quad \text{by (2).}$$

This means that there are no propositions at all within the rectangle of Fig. 1. But then there are no true propositions there either:

(6) $\sim (\exists p) . (f\,p \,.\, p).$

But (6) is another formulation of "$(p) . f\,p \supset \,\sim p$" which by (3) is equal to a. Therefore a is true. But if a is true, it must be a proposition, assuming in agreement with ordinary logic that nothing but a proposition can be true. Thus if *a* is not a proposition, it is a proposition and *vice versa*. This is again a circle in argument which is just as bad as the vacillation between (4) and (5). Thus Russell's treatment cannot resolve the paradox of Fig. I.*

* A reviewer of my version of the "Epimenides", C. H. Langford, decided that it does not differ "relevantly" from the original paradox (*The Journal of Symbolic Logic*, vol. 3, No. 1, 1938). His decision comes from a lack of understanding of my version as well as of the usual (i.e. Russell's) solution of the paradoxes. According to the reviewer, "It is not the case that the usual resolution of the paradox Ushenko cites is to the effect that the sentence within Fig. I is meaningless. . . . On the usual view, the sentence in question expresses a second-order proposition. . . ." But even if *a* could be reinterpreted as a second-order proposition, in its original interpretation, when it was intended (as it was in my exposition) to be about itself, it would have to be ruled out, on the usual view, as meaningless. However the important point, which the reviewer did not see, is that a reinterpretation of *a* as a second-order proposition is impossible, because a second-order proposition must be about first-order propositions, whereas *a* is not about first-order propositions; *a* is about propositions within the rectangle of Fig. I, and since, with the possible exception of *a* itself, there are no propositions there there are no first-order propositions there either. The reviewer should keep in mind Russell's own words: "It is important to observe that, since there are various types of propositions and functions . . . all phrases referring to 'all propositions' or 'all functions', or to 'some (undetermined) proposition' or 'some (undetermined) function' are *prima facie* meaningless, though in certain cases they are capable of an unobjectionable interpretation." (Op. cit., p. 166.) Professor Church has criticized my argument from a standpoint which is opposite to that of Mr. Langford, viz. he thinks that the statement *a* is meaningless because it contains the phrase "all propositions" without an indication of their order. But an indication of the order is not necessary if the phrase refers to a well-defined class of propositions, each of these being of a determinate order, as when we say that "All propositions on p. 1 of this book are either true or false". The sentence *a* is not about all propositions without restriction, but about "all-propositions-within-the-rectangle-of-Fig. I"; the hyphened expression may stand for a null class, but a null class is a well-defined class.

Nor would Ramsey's distinction of different "ways of meaning" be of any help here. For the expression *a* involves "writing" instead of "meaning", it is written within a rectangle and is itself about propositions written within a rectangle; and "writing", unlike "meaning", has no different senses which might lead to confusion or falsehood.*

A very simple resolution of this paradox would be a demonstration that it is not a paradox but a downright contradiction. Of course, the paradox leads to a contradiction when its implications are made explicit, but it remains a paradox so long as it appears that its formation agrees with the principles of logical significance while its descriptive import is an empirical fact as given by Fig. I. On the other hand, there would be no paradox if one could show that *a* violates the principles of logical formation and

* There might be a concealed reference to "meaning" in *a*, if as Professor Church pointed out to me, we must assume that only sentences and not propositions can be written and therefore take *a* as a sentence which *means* a proposition. This objection, it seems to me, is of an epistemological order. On purely logical grounds the question whether or not a proposition can be written remains open. So far as my own epistemological stand is concerned I believe that a proposition is "embodied" in a sentence, so that the former is written down along with the latter. Even if a proposition were a universal of which the corresponding sentences are instances, it would be present literally, as a whole, in each of its exemplifying instances.

because of this violation is a concealed expression of "$p \,.\, \sim p$". Accordingly, I suggest that *a* is not a paradox because it contradicts itself by contradicting the principle of logical significance which belongs to any proposition-form, viz., the "claim to truth". In its usual symbolic formulation this principle is given as an equivalence:

$$p \equiv (p = 1).$$

Should p assert its own falsehood, it would assert that "$p = 0$" and the equivalence would be transformed into:

$$p \equiv (0 = 1),$$

which is a contradiction. Thus we must reject *a*, which asserts its own falsehood, not as a paradox but as an outright contraction of the form "$0 = 1$" or "$p \,.\, \sim p$". Indeed, since *a* has an implicit claim to truth, it can be given explicitly as the conjunction:

a. the proposition expressed by "*a*" is true.

But in virtue of its meaning which is intended to apply to itself, *a* can be replaced in the conjunction by the weaker constituent:

The proposition expressed by "*a*" is false.

The result gives the contradiction:

The proposition expressed by "*a*" is false. The proposition expressed by ""*a* is true.

It is easy to show that the same principle can be applied to the original "Epimenides" and, in general, to any expression of the form:

All propositions of the kind, of which this is one, are false.

But I think a different treatment is required for the singular form:

(*a*) This proposition is false.

The difficulty here does not depend on self-attribution of falsehood, since a similar difficulty, and infinite regress, is found in the form:

(*b*) This proposition is true.

Hence the singular form (*a*) is not a mere variation of the "Epimenides", and must be referred to some different principle which could also solve the difficulty of (*b*). Such a new solution is found in the observation that (*a*) and (*b*), when intended to apply to themselves are not propositions but propositional functions, whose evaluation would always give falsehood.

Take the form (*a*). Since it is intended to be a statement about itself, it can be expressed as:

(*a*) is false,

where (*a*) is supposed to name the whole expression. But no sign can name (or single out) an object whose constitution involves that very sign, therefore (*a*) must be a pseudo-name, in fact it is nothing but a variable. And if (*a*) is a variable, then the expression "(*a*) is false" must be a propositional function whose substitution for (*a*) would be a statement about a proposition function, viz.:

"(*a*) is false" is false,

which is false or meaningless because a propositional function cannot be false. If one tried to avoid the use of a pseudo-name by interpreting "this" in "this proposition is false" not as a demonstrative symbol but as a description of "the proposition under consideration", one would miss directness of reference with a resulting ambiguity, and, which is worse, one would violate a distinction of significance by confusing a description of a proposition with the proposition itself. For to say that the phrase "the proposition (*a*)" means the same as "the proposition (*a*) is false" is to identify a description with a would-be proposition of which the former is a constituent. Some such confusion I find in Grelling's version given on p. 49, where "*q*" is first introduced not as a name of a proposition, but as an abbreviation of the definite description "the proposition on the 13th line of p. 49," and later identified with (1), although (1) is advanced as a proposition and not as a descriptive phrase.

But even if consistency alone were the right tool to work with the epistemological paradoxes, recourse to *significance* must be made when we come to Russell's "class of classes which are not members of themselves." Only I think that Russell is too summary when he asserts that all attributes indicate a structure. Let us try to discriminate. For example, we may compare the attributes "conceivable" and "yellow". Certainly "conceivable" means "an object of conception" and this phrase is a structure. But "yellow" can be visualized, and therefore thought of, in isolation from, i.e. without imagining it together with, a particular object. The evidence of

this example suggests a division of predicates into two groups: the predicates which are and the predicates which are not propositional functions. Of course, such a division if actually carried out, would depend on the connotation of predicates, but in abstraction from connotation it can be taken as a restrictive condition of logical formation to the effect that a predicate which is a structure or a propositional function cannot take another such predicate as a value of its argument. If $f(x)$ is a propositional function and g is a predicate without a structure, $f(g)$ is significant (unless further consideration of connotation rules it out), but $f(fx)$ is not. This condition excludes Russell's paradox. Let us take the attribute "Impredicable". This is a propositional function to which we have already attached the symbol:

$$I(x).$$

In conformity with the restrictive condition of significance the argument x of this function must take its values within the range of attributes which, like "yellow", have no structure. Hence "Impredicable" cannot be predicated of itself. The expressions "I (I)" and "$\sim$ I (I)" are illegitimate because they are short for the forms:

$$I(Ix); \quad \sim I(Ix),$$

which the restrictive condition rules out.

The paradoxes of logic have had a long and vexatious history, and I would not be surprised if my own resolutions will be found faulty. But the method of individual treatment always allows for a 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.