SUPPLEMENT II

THE THEORY OF TYPES[1]

BY PAUL WEISS[*]

It would seem from the interpretation that Whitehead and Russell put on the theory of types, that it is impossible or meaningless to state propositions which have an unrestricted possible range of values, or which, in any sense, are arguments to themselves. Thus on the acceptance of the principle that statements about all propositions are meaningless,[2] it would be illegitimate to say, "all propositions are representable by symbols," "all propositions involve judgment," "all propositions are elementary or not elementary," and if no statement could be made about all the members of a set,[3] it would be impossible to say, "all meanings are limited by a context", "all ideas are psychologically conditioned", "all significant assertions have grammatical structures", etc., all of which are intended to apply to themselves as well. The theory seems also to make ineffective a familiar form of refutation. General propositions are frequently denied because their enunciation or acknowledgment depends on the tacit supposition of the truth of a contradictory or contrary proposition. Such refutations assume that the general proposition should be capable of being an argument of the same type and to the same function as its own arguments, so that according to Whitehead and Russell, they fallaciously refute "by an argument which involves a vicious circle fallacy".[4]

That these limitations on the scope of assertions or on the validity of refutations are rarely heeded is apparent even from a cursory examination of philosophical writings since 1910. Thus Russell, apropos to Bergson's attempt to state a formula for the comic says,[5] "it would seem to be impossible to find any such formula as M. Bergson seeks. Every formula treats what is living as if it were mechanical, and is therefore by his own rules a fitting object of laughter." The characterisation of all formulae, even though it refers to a totality, seems to Mr. Russell to be of the same type as the formulae characterised.

[1] Chap. II., *Principia Mathematica*

[*] [*Reprinted from* MIND: *a Quarterly Review of Psychology and Philosophy. Vol. XXXVII., N.S., No.* 147; with minor corrections.]

[2] P. *37, ibid.* (second edition).

[3] P. 37, *ibid*

[4] P. 38, *ibid.*

[5] "Prof. Guide to Laughter," *Cambridge Review,* Vol. 32, 1912, and Jourdain's *Philosophy of Mr. B*tr*nd R*ss*ll,* pp. 86-7

If the theory were without any embarrassments of its own, and were indispensable for the resolution of the so-called paradoxes[1] (which no one seems to believe), there would be nothing to do but to acknowledge the impossibility of cosmic formulations, as well as the inadequacy of philosophic criticisms, and to pass charitably over such remarks as Russell's as mere accidents in a busy life. However, the statement of the theory itself involves the following difficulties in connection with (1) its scope, (2) its applicability to propositions made about it, and (3) its description.

1. It is either about all propositions or it is not.

A. If it were about all propositions it would violate the theory of types and be meaningless or self-contradictory.

B. If it were not about all propositions, it would not be universally applicable. To state it, its limitations of application would have to be specified. One cannot say that there is a different theory of types for each order of the hierarchy, for the proposition about the hierarchy introduces the difficulty over again.

2. Propositions about the theory of types (such as the present ones, as well as those in the *Principia*) are subject to the theory of types, or they are not.

A. If they were, the theory would include within its own scope propositions of a higher order, and thus be an argument to what is an argument to it.[2]

B. If they were not, there would be an unlimited number of propositions, not subject to the theory, that could be made directly or indirectly about it. Among these propositions there might be some which refer to a totality and involve functions which have arguments presupposing the function.

3. The statement of the theory of types is either a proposition or a propositional function, neither or both.

A. If it were a proposition, it would be either elementary, first order, general, etc., have a definite place in a hierarchy and refer only to those propositions which are of a lower order. If it were held to be a proposition of the last order, then the number of orders would have a last term, and there could not be meaningful propositions made about the theory. The *Principia* should not be able to say, on that basis, just what the purpose, character and application of the theory is.

B. Similarly, if it were a propositional function, it would have a definite place in a hierarchy, being derived from a proposition by generalisa-

[1] Paradoxes, though contrary to common opinion, may be and frequently are true. Paranoumena, violating principles of logic or reason, if they are not meaningless, are false, and it is only they which are capable of logical analysis and resolution. What the *Principia* attempts to do is to solve apparent paranoumena with a real paranoumenon.

[2] P. 39, *Principia Mathematica*

tion. It could not refer to all propositions or propositional functions, but only to those of a lower order.

C. If it were neither it could not be true or false, nor refer to anything that was true or false. It could not apply to propositions, for only propositions or propositional functions, in a logic, refer to propositions.

D. If both at once, it would be necessarily self-reflexive.

a. If as function it had itself as value, it would refer to itself. But the theory of types denies that a function can have itself as value.

b. If as function it had something else as value, it would conform to the theory, which insists that functions have something else as values. The theory then applies to itself and is self-reflexive, and thus does not apply to itself. As, by hypothesis, it is a value of some other function, there must be propositions of a higher order and wider range than the theory of types.

It is no wonder that the perpetrators of the theory have not been altogether happy about it ! What is sound in it—and there is much that is—is best discovered by forgetting their statements altogether, and by endeavouring to analyse the problems it was designed to answer, without recourse to their machinery. The result will be an acknowledgment of a theory of types having a limited application, and a formulation of a principle which will permit certain kinds of unrestricted general propositions.

To do this we shall deal in detail with two apparent paranoumena dealt with in the *Principia*, where the difficulty is largely *methodological*. We shall then treat of Weyl's "heterological-autological" problem, where the difficulty is due to a confusion in *meanings*. Those problems which cannot be dealt with under either heading will be those which need a theory of types for their resolution.

1. *Epimenides*. The proposition "All Cretans are liars" must be false if it applies to Epimenides as well, for it cannot be true, and only as false has it meaning. If it were true, it would involve its own falsity. When taken as false, no contradiction, or even paradox, is involved, for the truth would then be "*some* Cretans tell the truth". (The truth could not be "all Cretans tell the truth" for Epimenides must be a liar for that to be true and by that token it must be false). Epimenides himself would be one of the Iying Cretans, and one of the lies that the Cretans were to make would be "all Cretans are liars". Thus if Epimenides meant to include all his own remarks within the scope of the assertion, he would contradict himself or state a falsehood. If it be denied that a contradictory assertion can have meaning, he must be saying something false if he is saying anything significant. Had he meant to refer to all other Cretans there is, of course, no difficulty, for he then invokes a kind of theory of types by which he makes a remark not intended to apply to himself. All difficulty disappears when it is recognised that the formal implication, "all Cretanic statements are lies" can as a particular statement be taken as one of the values of the terms of this implication. Letting *Ep ! p* represent "Epimenides once asserted p"; Φ represent "Cretanic" and *p* represent a statement or

proposition. then for "All Cretanic statements are false (or lies)," we have:

1. $\Phi p \, . \supset_p . \sim p.$

And as Epimenides is a Cretan, for any assertion he makes we have:

2. $Ep \, ! \, p \, . \supset_p . \, \Phi p.$

As No. 1 is an argument to the above—it being Epimenides' present remark— we get:

3. $Ep \, ! \, \{\Phi p \, . \supset_p . \sim p\} \, . \supset . \, \Phi\{\Phi p \, . \supset_p . \sim p\}$

No. 1, as a Cretanic statement, is an argument to No. 1 as a formal implication or principle about Cretanic statements, so that:

3A. $\Phi\{\Phi p \, . \supset_p . \sim p\} \, . \supset . \sim \{\Phi p \, . \supset_p . \sim p\}$

No. 3 and No. 3A by the syllogism yield:

3B. $Ep \, ! \, \{\Phi p \, . \supset_p . \sim p\} \, . \supset . \sim \{\Phi p \, . \supset_p . \sim p\}$

so that in this instance Epimenides lied.

It is important to note that No. 1 states a formal implication, and that No. 3, No. 3A and No. 3B employ No. 1 as a particular assertion or specific argument to their functions. No. 3A is an instance of the implication expressed by No. 1, and is this instance because of the particular argument it does have. It states the fact that " 'all Cretanic statements are false' is a Cretanic Statement," implies that " 'all Cretanic statements are false' is false". Substitution of another argument would give a different instance; though of course of the same implication. The implication contained in its argument does not have instances. " 'Some Cretanic statements are false' is a Cretanic statement" or " 'This Cretanic statement is false' is a Cretanic statement" are not instances of " 'All Cretanic statements are false' is a Cretanic statement," but of "P is a Cretanic statement". These three propositions have different subjects; they are different values of the same propositional function. That these subjects have relations to one another is of no moment. "My wife loves me" and "my mother-in-law is old (or loves me)" are two distinct and logically independent propositions, even though there is a relationship between the two subjects.

It is because any considered general proposition is at once an individual fact, and a formal implication or principle, with many possible arguments, that it is capable of being taken as an argument to itself. All propositions about words, logic, truth, meaning, ideas, etc., take arguments which fall in these same categories, and in so far as such a general proposition is stated in words, determined by logic, etc., it should, as such a fact, be an argument to itself as a formal implication. The principle must be false if this cannot be done, for it is sufficient, in order to overthrow a proposition of this kind, to produce one argument for which it does not hold. One may limit the principle by asserting that it holds for "all but . . .", in which case it is a *restricted* general proposition. Nominalism, association of ideas, scepticism, the theory of universal tautology, the denial of logic are defended in propositions which cannot take themselves as arguments, and which as facts are arguments to contradictory principles. Their contradictory principles therefore hold sometimes at least, so that these doctrines must be false if they are put forward without restriction, and cannot be universally true, if, in Bradley's words, they "appear".

2. "I am lying"—if it be taken in isolation from all fact—is a meaningless statement. There must be some objective truth that is distorted, and unless it is provided the assertion has no significance. This proposition means either, "I am lying about X"; "I always lie," or "I have always lied". The first can be either true or false without giving rise to any problem, except where "all my assertions" is made an argument to X, in which case it is equivalent to either the second or third formulation. "I always lie" involves the same situation as with Epimenides, and the proposition is false. The supposition of its truth would involve a contradiction; the supposition of its falsity means simply that I sometimes lie and sometimes tell the truth. If what is meant is that "I have always lied" that does not involve a contradiction, for what is intended is a restricted proposition, applying to *all but* the present one. It can be true because it does not apply to all propositions; if it were false, then sometimes I lied and sometimes I did not. In short, there is nothing like a self-reflective universal liar, which is an interesting moral conclusion to derive from a logical analysis. Similarly, there cannot be a thorough scepticism held by the sceptic to be valid.

Prof. Whitehead (to whom I am also indebted for the notation) has pointed out to me that wherever a conjunction of propositions results in a *reductio ad absurdum*, there is no way of determining on logical grounds alone which of the antecedents fails, or is false (though one at least must be). Thus in the case of Epimenides we have:

$$4.\ \underbrace{\{\Phi p\,.\supset_p.\sim p\}}_{(A)}\,.\,\underbrace{\{Ep\,!\,p\,.\supset_p.\,\Phi\, p\}}_{(B)}\,.\,\underbrace{Ep\,!\,\{\Phi p\,.\supset_p.\sim p\}}_{(C)}\,.\supset.\sim\underbrace{\{\Phi p\,.\supset_p.\sim p\}}_{(D)}$$

It is because B and C are in that case assumed to hold, that we can say that A must fail. If the truth of all these antecedents were undetermined, we should have merely the general rule: a *reductio ad absurdum* has as a necessary condition the conjunction of one or more false propositions. Transposition—

$$4'.\ \underbrace{\{\Phi p\,.\supset_p.\sim p\}}_{(D)}\,.\supset.\sim\underbrace{\{\Phi p\,.\supset_p.\sim p\}}_{(A)}\,.\vee.\,\underbrace{\{Ep\,!\,p\,.\supset_p.\,\Phi\, p\}}_{(B)}\,.\vee.\sim\underbrace{Ep\,!\,\{\Phi p\,.\supset_p.\sim p\}}_{(C)}$$

makes it apparent that to deny the conclusion of a *reductio ad absurdum* is to imply that at least one of the antecedents is false.

In connection with the *reductio ad absurdum* involved in the assertions, "I always lie" and "I always doubt," No. 4B reduces to the tautologies: "If I assert *p*, *p* is my assertion," and "If I doubt, the doubt is mine". In these cases, the only alternatives left are the denial of the fact of the assertion (No. 4C), or the truth of the principle itself (No. 4A).

3. Weyl's heterological-autological contradiction[1] is the result of a material fallacy of amphiboly in connection with the employment of adjectives. The simplest form of such a fallacy is due to a failure to distinguish between an adjective as substantive and an adjective as attribute. Thus if we treat both the subject and attribute in "large is small" and "small is large" as attributes united by a copula expressing identity (instead of reading it as "large is a small word", "small is a large word") we could say "whatever is small is large, and whatever is large is small". No one, I believe, since the Megarics, has been troubled by this particular confusion.

The present problem is the result of a confusion, not between substantive and adjective, but between an adjective which expresses a property, and an adjective which expresses a relation between this property and the substantive. All words can be described in terms of a property—they are long, short, beautiful, melodious, etc., words. They can be classified in accordance with these properties, giving us the class of long words, short words, etc. They can also be classified as either "autological" or "heterological," depending on whether or not the same word is at once substantive and property-adjective; the terms "autological" and "heterological" expressing relationships between the substantive and adjective.

The autological class is made up of words, each of which expresses a property which it possesses; though all of them have unique properties. If "short" be short, and if "melodious" be melodious, they would both be members of the autological class; though in addition, "short" would be a member of the class of short words, and "melodious" would be a member of the class of melodious words.

The heterological class is made up of words, each of which expresses a property which it does not possess. If "long" be short, and if "fat" be thin, they would both be members of the heterological class; although here also "long" would be a member of the class of short words, and "fat" would be a member of the class of thin words. Though when classified according to the relationship of the adjective to the substantive, "short" would be an autological word and "long" a heterological word, they would both be members of that class which was defined in terms of the properties of words—being in this case, members of the class of short words.

Now if heterologicality were a property that a word could have, and if the word "heterological" had that property, it would be a member of the autological class, for it would then possess a property that it expressed. But it would also be a member of a class of words which had the *property* of heterologicality. This class is determined by taking the properties of words, and if it be called

[1] Briefly stated it is: all words which express a property they possess are autological; all words which express a property they do not possess are heterological. If 'heterological' is heterological it expresses a property it possesses-and is thus autological; if it is autological, it expresses a property it does not possess and is therefore heterological. *Das Kontinuum,* p. 2.

“heterological”, must be distinguished from that class which was determined not by properties, but by the relationship between properties and substantives.

If there were a property like autologicality and if “heterological” had that property,[1] it would be a member of the heterological class, for it would express a property which it did not possess. But it would also be a member of the class of words which possessed autologicality and could thus be classified.

Thus if “heterological” had the property of autologicality, it would be in the heterological class owing to the *relation* which held between the property and substantive (or between a property it possessed and the property it expressed); but it would be in the class of autological words, owing to a *property* it possessed. If it had the property of heterologicality, it would be in the autological class on the basis of the *relation*, and in the class of heterological words on the basis of *property* classification. There is no difficulty in considering something as a member of two distinct classes, owing to the employment of different methods of classification. There is no contradiction in saying: “ ‘heterological’ expresses the property heterologicality, possesses the property autologicality, and the relation between these properties is heterological, or that it expresses and possesses the property heterologicality and the relation between them is autological.” Similarly, Richard’s contradiction, Berry’s contradiction, and that involving the least indefinable ordinal, are resolvable by recognising that “nameable” and “indefinable” are used in two sharply distinguishable senses. They do not require a hierarchy, but a discrimination in the methods of description.

When a distinction is made between a class and its membership (the distinction between a number of numbers and a number is a particular case of this), and between a relation of objects and a relation of relations, the requirements for the solution of the other mathematical problems are provided. A class is other than its members, and a relation, like all universals, transcends any given instance or totality of instances. As they have characters of their own, universals can be described in terms of other universals, which in turn transcend them. Arguments are of a different “type” than functions, just so far as they have different logical characteristics, *i.e.* are different kinds of logical facts. The class which is an argument to a function about classes has, as argument, a different logical import than the function, and its arguments have a different import from it. This is true of all functions, restricted and unrestricted alike, for it means simply that they are discriminable from their arguments. They can, despite this difference, have characteristics in common with their arguments, and are to that extent unrestricted. Thus in the case of “the class of those classes which are identical with themselves,” the class of classes can be

---

[1] ‘Heterological,’ in fact, has the properties of being long, polysyllabic, etc., and it is questionable whether there are properties like autologicality and heterologicality possessed by words. If there be no such properties, ”heterological” is a member of the class of long words, polysyllabic words, etc. In addition it would be one of the terms related by the heterological relation, which fact would not make it have the *property* of heterologicality.

taken simply as a class, without logical embarrassment. Yet a class of classes differs from a class, and must therefore be capable of a different characterisation, and thus also be an argument to a function of a different type. With some classes, it may not be possible to consider them as arguments to their own functions, without uncovering a contradiction. In such cases (*e.g.* the class of those classes which are not members of themselves, and the relations which are connected by their contradictories), it is the difference between the function and the argument that is of moment. That *some* functions cannot take themselves as arguments does not indicate that *all* functions are restricted in scope, but simply that they are *non-restricted.* Some classes and functions are restricted and some are not. To say that all are restricted because some are is an obvious fallacy.

Whenever, as individual, a general proposition is in the class of those objects of which it treats, but cannot be considered as an argument to itself, it is either false or restricted in scope. If the second, its range of arguments must be specified. Accordingly, we can state as a *necessary* condition for the truth of a general proposition, whose scope is unspecified, that when it has a character, which is one of the characters about which it speaks, it *must* be an argument to itself. Thus if Bergson adequately described the comic, his formula should be an object of laughter, and if the theory of types is universal in application, it should be capable of being subject to itself. Conformity to this condition indicates that the unrestricted proposition is *possibly* true; not that it is necessarily true. To demonstrate that such a proposition was necessarily true, it would be essential to show that the supposition of its falsity assumes its truth. That there is danger in applying this rule can be seen from the consideration of some such proposition as: “Everything is made up of language elements”. Its denial will be made up of language elements, and would seem to demonstrate that the proposition was necessarily true. Supposition of the falsity of a proposition, however, means verbal denial only in so far as the proposition applies to the realm of language. If it applies to everything, supposition of its falsity involves the positing of the objects of assertions; not the assertions. A necessary unrestricted proposition about everything can be supported only by a demonstration that the supposition of an argument for which it does not hold is self-contradictory. If the proposition has to do with grammar, meaning, logic, judgment, etc., the conditions for a necessarily true and unrestricted proposition would be: 1. the assertion of it is an argument to it; 2. any possible denial is an argument to it. That “any possible denial” rather than “any given denial” is required, is apparent from the consideration of the following propositions: “All sentences are made up of eight words,” “No sentence is made up of eight words”. Each of these contains eight words. It is because of the fact that we can formulate propositions such as, “It is false that every proposition must be made up of eight words, ” that the condition is seen not to have been met.

An unrestricted proposition applies to every member of the category, and has some aspect of itself as value. It is in some sense then a determinate in the category which it determines. If the proposition refers to some other category than the one to which it as fact, or some aspect of it as fact, belongs, it is restricted. Thus "all men are mortal" is neither man nor mortal, and as condition does not determine itself as fact. Any proposition referring to that statement would be of a different type, and would deal with its truth, falsity, constituents, historical place, logical structure, etc. Though the unrestricted propositions have no limitations, the category to which they refer may have. Epimenides' remark, for example, referred only to Cretans. As his assertion was a determinate in the category, and as his statement of the supposed conditions imposed on the members of that category was not a possible argument to the general proposition, the general proposition was seen to be false or restricted. Had he said, "All Cretans tell the truth", he would have stated an unrestricted proposition which was possibly true. It could not be said to be necessarily true unless Cretans and lie, against the evidence of history, were actually contradictories.

Accordingly, we shall say: *All true unrestricted propositions are arguments to themselves; or by transposition, those propositions which are not arguments to themselves are either restricted or false.* As this proposition can take itself as argument it is possibly true. Unless no proposition is possible which does not conform to it, it cannot be said to be necessarily true. I have not been able to demonstrate this and therefore accept it as a definition or "methodological principle of validation". The theory of types, in its most general form, may be stated as: A *proposition or function of order n, which cannot be an argument to itself, is, as fact, an argument of a proposition or function of order* $n+1$.

In accordance with the scheme of the criticism of the theory of types, we can describe our principle as (1) applying to all propositions, including (2) those which refer to it. (3) It is a formal implication with itself as one of its arguments. The theory of types, on the other hand, (1) does not apply to all propositions, but only to those which are restricted, (2) *may* apply to those propositions which refer to it, and (3) is a formal implication which cannot take itself as argument.

The theory of types cannot be an unrestricted proposition about all restricted propositions. As an unrestricted proposition it must take itself as argument; but its arguments are only those propositions which are not arguments to themselves. It cannot therefore be unrestricted without being restricted. Nor can it be a restricted proposition about all restricted propositions for it would then be one of the restricted propositions, and would have to take itself as argument—in which case it would be unrestricted. Hence it cannot be restricted without being unrestricted. Three possible solutions may be advanced. The first is that the theory of types is restricted and does not apply to *all* restricted propositions, but only to *some* of them. It is not an argument to itself but to some other proposition about restricted propositions. This in turn will have to be restricted and refer only to some propositions, and so on, giving us theories of types of various orders. The proposition made about the totality of these orders would be of a still higher order and would in turn presuppose a higher order *ad infinitum.* The theory of types thus depends on theories of types of theories of types without end. This seems probable on the ground that the theory is based on the recognition that no proposition can be made about all restricted propositions, so that it must by that very fact admit that it cannot apply to all of them. Instead, therefore, of the theory of types applying to all propositions, and determining them in various orders, it does not even apply to all of a given class of them. This interpretation would not affect unrestricted propositions, and would merely show that the determination of restricted propositions is subject to determinations without end.

The second possibility is suggested by the consideration of a proposition such as: “all truths are but partially true”. If that were absolutely true, it would contradict itself, and if it were not, could apply only to some truths. Considered as referring to the necessary limitations which any finite statement must have, it would take itself as argument in so far as it was finite, thus indicating that it was absolutely true about finite propositions, and yet not absolutely true as regards all truths. By pointing out the limitations of a finite statement it indicates that there is an absolute truth in terms of which it is relatively true. On this interpretation, any condition which imposes universal limitations is unlimited in terms of what it limits, but limited in turn by some other condition. One might hold, therefore, that the theory would be unrestricted as regards restricted propositions, and restricted as regards all propositions, and would point to a higher principle which limits it.

The third possibility is to allow for “intensive” propositions which are neither restricted nor unrestricted, being incapable of any arguments. The theory of types could be viewed as such an intensive proposition, and what we have called its arguments, would merely “conform” to it. This interpretation means the downfall of a completely extensional logic, and a determination of an extensional logic as subordinate to an intensional one.

There are difficulties in each of these interpretations. The last seems to me to be best. In any of these cases, however, a restricted proposition which refers to some other than the restricted aspect of the theory would be subject to the theory and the principle we have laid down about unrestricted propositions could still hold. Those restricted propositions which refer to the restricted character of the theory would not be an argument to it on the first, would be an argument to it on the second, and would neither be nor not be an argument to it on the third solution.

To briefly summarise: The theory of types must be limited in application. Not all the problems it was designed to answer require it; another principle of greater logical import is desirable; while for the resolution of the problems in which it is itself involved, very drastic remedies are necessary. No matter how the theory fares, the possibility of the methodological principle and the possibility of other solutions for the so-called paradoxes, indicate that it is at least not as significant an instrument as it was originally thought to be.