Confusion of Symbols and Confusion of Logical Types
Author(s): Susanne K. Langer
Source: *Mind*, New Series, Vol. 35, No. 138 (Apr., 1926), pp. 222-229
Published by: Oxford University Press on behalf of the Mind Association
Stable URL: http://www.jstor.org/stable/2249354
Accessed: 20/06/2014 17:44

V.—DISCUSSIONS.

CONFUSION OF SYMBOLS AND CONFUSION OF LOGICAL TYPES.

THE ambiguities of language have ever been responsible for a host of conflicting notions in logic and philosophy. These ambiguities were expected to disappear, once and for all, before the conquering power of symbolic expression, which should fix unalterably the only and entire meaning of every concept employed in the discourse. The rich and irrelevant connotations of words, accumulated through centuries of common usage, were supposed to be banished by the introduction of precise, novel, specialised symbols. But the use of symbolism has an unforseen pitfall of its own, the exact converse of the danger that lurks in ordinary language; our symbols are liable to confusion by their very abstractness. In a complicated logical structure it may become difficult to observe that $x = x$, just because "x" is so non-connotative that it is hard to remember its ear-marks once they are fixed. The ambiguity of the variable is apt to suggest, falsely, an ambiguity in its formal relations, and this fallacy is productive of interesting-looking absurdities. Such are, for instance, the famous paradoxes of logic, which Bertrand Russell seeks to obviate by the theory of types.[1]

Oftentimes the hardest part of the solution of a problem lies in stating the difficulty. This is true in the present case. Paradoxes, antinomies and other sophisms have been with us since the beginning of philosophy, because no one could discover the confusion of concepts which engendered them. To Mr. Russell belongs the credit for this discovery; and since he has revealed the source of fallacies, their elimination from logical and mathematical theory can only be a matter of time. It is true that the theory of types is not entirely satisfactory, but now that he has shown the way, it may be the task of later logicians, coming in the wake of his pioneering, to solve the problems he has articulated. In the recent new edition of *Principia Mathematica* we already find references to two alternative solutions, that of Dr. Chwistek and that of Mr. Wittgenstein, the latter given in considerable detail. Mr. Wittgenstein accepts the theory of types, but would substitute for the logical "axiom of reducibility" another, somewhat more philosophical

[1] We have Prof. Whitehead's authority to state that Mr. Russell is the originator of the type-theory, and sole author of the new Introduction to the second edition of *Principia Mathematica*.

assumption, which we might call the "axiom of truth-functions." Finally, Mr. Russell himself has presented us with an alternative, the assumption that a function can only appear in a logical matrix through its values.[1] Yet it seems curious that, in order to decide whether a proposition is or is not self-contradictory, we must resort to a *rule* of any kind. And it is my purpose here to show that, given the non-formal concepts of *Principia*—the "interpretational" elements of *proposition, function, truth,* etc., that is to say: the concepts in terms of which the formal calculus is there interpreted,—the "assumption" Mr. Russell makes in the new Introduction follows as a theorem, and is validated by the nature of his entities.

The first step toward locating the source of confusion was the discovery that most of the paradoxes in question—in his opinion, all—were interpretations of one sort of construct: that they could all be expressed by a function of the form, $\phi(\phi\hat{x})$. But although the paradoxical propositions have been avoided by ruling out all such "reflexive" functions, we are faced with a new difficulty; for it seems upon superficial inspection as though not only nonsensical, but also many apparently valid assertions had been banished by the type-theory, and this gives the solution an arbitrary air. So far the form $\phi(\phi\hat{x})$ has been really "ruled" out, not because it is intrinsically untenable, but because it is untrustworthy—we have to dispense with *all* propositions of this form because we have empirical evidence that *some* of them are "vicious." If, however, the fallacy is logical at all, and belongs to the *structure* $\phi(\phi\hat{x})$, then it should be possible to show, without constructing a theory or accepting a new axiom, that the form $\phi(\phi\hat{x})$ is really patent nonsense and no sensible interpretations of it are possible. The theory of types is a necessary safeguard if we are to employ the calculus of *Principia* as a "mathematics without meaning," for then we require a rule to govern the manipulation of every sort of mark, and the fact that $\phi(\phi\hat{x})$ cannot occur must be explicitly stated in a system where nothing is to be inferred from the *meaning* of the marks. But then we need no theory; the philosophical grounds for our ruling do not concern the manipulator of the marks. It is otherwise if we treat the marks as *symbols*, as Mr. Russell does; then the rule "$\phi(\phi\hat{x})$ must be meaningless" expresses merely that the concepts corresponding to ϕ and to $\hat{x}$ cannot be combined in any way analogous to the pattern, $\phi(\phi\hat{x})$. And if this is a fact, then a consideration of those concepts will make it evident, so that the prescriptive statement: "$\phi(\phi\hat{x})$ must be meaningless" is superfluous, and may be displaced by the *descriptive* proposition: "$\phi(\phi\hat{x})$ *is* meaningless."

There is certainly an ambiguity in the expression $\phi(\phi\hat{x})$; Mr. Russell has sought to locate that ambiguity in the function. ϕ, he claims, has more than one meaning, and the ϕ outside the bracket is not really the same as the ϕ within. The two have different

[1] Introd., p. xxix.

ranges of significance; in the propositions "x is false" and "'x is false' is false," the word "false" has two different meanings respectively. Now it is hard to convince ourselves that we really do not mean by "false" No. 1 what we mean by "false" No. 2; the systematic ambiguity of truth and falsehood is certainly not obvious to common sense, and it is only by reduction of the proposition to the standard form, and the application of the vicious-circle principle, that any statement may be known to contain similar confusion of meanings. A list of such "ambiguous" concepts is given in *Principia Mathematica* (p. 64);[1] but there seems to be no method, except the chance discovery of paradoxes, for the recognition of such concepts.

There is, however, another sort of ambiguity involved in propositional functions of the form $\phi(\phi\hat{x})$, and this may be shown to hold for any interpretation, and to be discoverable at sight, without looking to its consequences, because it is demonstrable *in abstracto*. This ambiguity lies in the argument, not in the function. It is the symbol $\hat{x}$ which suffers a change of meaning, not the ϕ; and this fact can be exhibited.

Mr. Russell maintains that all the traditional paradoxes of logic are instances of some function, $\phi\hat{x}$, with $\phi\hat{x}$ for its argument. It seems, however, that not only $\phi(\phi\hat{x})$, but also $\phi(\phi x)$ may occasionally give rise to paradoxes. In *Principia* $\phi\hat{x}$ is treated as though it were simply a generalisation of ϕx, so that the two fallacies seem to be instances of the same principle. But they are, in fact, of different origin and importance. For such a construct as $\phi(\phi\hat{x})$ does violence to the meaning of $\phi\hat{x}$, whereas $\phi(\phi x)$ involves only a misuse of the variable. The chief interest attaching to the latter case is that it is due to this sort of proposition that the theory of types is accused of ruling out valid as well as invalid situations, as we presently shall see.

The meaning of $\phi\hat{x}$ is a very subtle affair, and our appreciation of it is rendered more difficult by a rather unfortunate choice of symbolism and terminology. Mr. Russell calls $\phi\hat{x}$ a function, and its appearance, in the guise of a Greek letter followed by a small Latin letter, certainly suggests the same conception as, for example, fa, fx, ϕx. These latter three are functions involving three different *degrees of ambiguity*. fa is one possible value for ϕa; ϕa, in turn, is a value for ϕx, that is to say, the x of fx ambiguously denotes the a in fa, and the ϕ of ϕx ambiguously denotes the f of fx or fa. On page 40 of *Principia*, Mr. Russell tells us that $\phi\hat{x}$ is a function whose values are ϕx, ϕy, ϕz, etc. This makes $\phi\hat{x}$ look like the same sort of construct as ϕx, ϕy, ϕz, . . . possessing simply a higher degree of ambiguity. In the original text we are given no further definition of $\phi\hat{x}$. But this defect is made up in the new Introduction, where we read: "We may define a function $\phi\hat{x}$ as that kind of similarity between pro-

[1] All references are to the 2nd ed.

positions which exists when one results from the other by the substitution of one individual for another."[1] (Perhaps we might define $\phi\hat{x}$ still more accurately as that structural element, common to certain propositions, whereon rests that kind of similarity . . .)

Now this definition makes clear the fact that the $\hat{x}$ in $\phi\hat{x}$ is not a variable, an ambiguous expression of x, y, z, etc. The $\hat{x}$, indeed, *is not an argument*, and *$\phi\hat{x}$ is not a genuine propositional function.* That which several propositions have in common is not something variable, not something ambiguous, and ambiguity is the essential characteristic of functions.[2] $\phi\hat{x}$ does not denote; it is not an expression of, but *a true abstraction from*, ϕx, ϕy, ϕz, etc. It might be called a "function-form"—a sort of astral-body of a function—where $\hat{x}$ is not an argument, but the logical place of an argument, which Frege has aptly but untranslatably called an "Argumentsstelle".

$\phi\hat{x}$ is an abstraction from ϕx, ϕy, etc.,—their *ϕ-ness*, so to speak. The $\hat{x}$ is here a very misleading symbol, because it looks like an argument to ϕ. It is really no such thing. $\phi\hat{x}$, or $\phi(\ )$ as I prefer to write it, is a symbol expressing a character of ϕx, ϕy, etc., or of their respective values. It means approximately, "the ϕ-ness of".[3] To distinguish it from a genuine propositional function, which ambiguously denotes a number of propositions, I shall call the concept $\phi(\ )$ an *abstractive function.* Its definition, as we saw above, presupposes the notion of genuine propositions or propositional functions. Therefore to write ϕx with the argument $\phi(\ )$, is to treat the argument to a propositional function as an abstraction from *that function with its argument.* A true abstraction, such as the purely formal structure of any entity, can be

[1] P. xxx.

[2] *Princ. Math.*, p. 39: "the essential characteristic of a function is *ambiguity.*"

[3] The precise meaning of $\phi\hat{x}$ is hard to render in any verbal form. It has been translated in participial and infinitive phrases, but this treatment is not satisfactory. If ϕx is taken to mean "x is a cat," then $\phi\hat{x}$ by the above interpretation would mean "being a cat". But "being a cat" does not express the cat-concept in the *form* by virtue of which it determines the class of cats. It is practically a substantive notion. Neither may we regard $\phi\hat{x}$ as the predicate "is a cat"; that meaning belongs to the ϕ in ϕx. "Is a cat" does not define a class, because it can take only one argument at a time (this may be an individual or group of individuals). $\phi\hat{x}$ seems rather to correspond to a notion for which the structure of English grammar does not exactly provide, and its nearest rendering would be some such barbarism as: "The being-a-cat-*ness* of". This renders approximately the ϕ in $\phi\hat{x}$; the $\hat{x}$ in turn, indicates the need of some argument and makes the function "incomplete"—essentially undetermined apart from its values, which have arguments. (Thus whatever difference there may be between the two ϕ's in $\phi(\phi\hat{x})$, is grammatical, not essential: *e.g.* if ϕx means "x is true," the difference would be between "is true," and "truth" or "being true" or what not, but not between sorts of truth.) I think the above interpretation is borne out by the definition in the new Introduction, quoted above.

shown only in some entity which "has" the structure. (This is the basis of Mr. Wittgenstein's mysticism.) Even though we may seek to make the structure more evident by varying the specific entities, as when we put x, y, z, for a, b, c, this is merely a psychological device to secure recognition of the similarity between the two sets of symbols, which arises from the same abstractive function expressed in both. We cannot in turn symbolise the function without simply employing another set of marks, say α, β, γ, which expresses the same abstractive function in exactly the same way. Therefore Mr. Russell's definition of a function, $\phi\hat{x}$, implies the proposition that a function can appear in a matrix only through its values. Instead of a hierarchy of types, pragmatically introduced, we have a hierarchy of abstractions, logically introduced, *i.e.* implied by the primitive notions.

The typical case where $\phi(\ )$ is supposed to figure as argument to a function *of the form* $\phi(\ )$, that is: to ϕx, where $\phi(\ )$ is abstracted from ϕx, ϕy, ϕz, etc., is the case where $\phi(\ )$ is treated as the definition of a class. (Note that, according to the new Introduction to *Principia*, there is no difference between classes and functions,[1] so that the symbol $\hat{x}(\phi x)$ is equivalent to $\phi\hat{x}$, and the separate assumption "that classes exist," which Mr. Russell refused to make,[2] becomes superfluous if we assume the existence of functions—a primitive notion without which *Principia* would hardly be comprehensible. I shall often employ the symbolism $\hat{x}(\phi x)$, because it is more familiar, but it should be borne in mind that this may also be read $\phi\hat{x}$.)

Our theorem is, then, that $\hat{x}(\phi x)$ cannot be a value for a function of the form $\phi(\ )$, because $\hat{x}(\phi x)$ as argument to such a function is ambiguous.

Dem.:

$\phi\hat{x} \equiv \hat{x}(\phi x)$

[p. xxxix]

$\phi\hat{x}$ is abstracted from *all* propositions ϕx, ϕy, ϕz, etc. Therefore it is an abstraction from ϕx.

[Def. of $\phi\hat{x}$, p. xxx].

Therefore x in ϕx denotes a part of a proposition (the argument) from which $\phi(\ )$ is derived. Thus the denotation of x stands in a fixed relation to $\phi(\ )$. If then x is taken to denote $\phi(\ )$, x denotes both $\phi(\ )$ and another concept to which $\phi(\ )$ stands in the assymmetrical relation of abstraction. Hence $x_1 \neq x_2$, which violates the canons of notation. The ambiguity lies in x, not in ϕ. Q.E.D.

In the light of this structural analysis, let us look once more at the paradoxes, ancient and modern, cited in *Principia* as examples of the vicious-circle fallacy, and enquire whether such a dual personality of x, rather than of ϕ, is not involved in every case of $\phi(\phi x)$.

"(1) The oldest contradiction of the kind in question is the

[1] *Princ. Math.*, p. xxxix. [2] *Ibid.*, p. 58.

Epimenides. Epimenides the Cretan said that all Cretans were liars, and all other statements made by Cretans were certainly lies. Was this a lie? The simplest form of this contradiction is afforded by the man who says 'I am lying'; if he is lying, he is telling the truth, and *vice versa*." [1]

The "simplest form" is, as a matter of fact, *not* the same sort of contradiction as that asserted by the unpatriotic Cretan. It is of the form $\phi(\phi x)$, and I reserve it for special consideration. But the statement of Epimenides, having for its subject *all* his statements, is truly of the form $\phi(\phi\hat{x})$. Now all his statements compose $\phi\hat{x}$; so that, if x is the argument to one of these, and also is $\phi\hat{x}$, then we have lost track of one argument, namely the one belonging to ϕx in virtue of which $\phi\hat{x}$ could be abstracted from it. It must have been a genuine proposition in order to give rise to $\phi\hat{x}$, and as such must have had an argument. But if we write the proposition ϕx and then let x denote both the argument in a proposition from which ϕx could be derived, and $\phi\hat{x}$, we simply give x a double meaning. We use it to denote both the extension of the function and something constitutive of that extension.

So we see that the testimony of Epimenides is without form, and therefore void. A statement about all statements is a structural impossibility, not because two different assertions are made about the same thing, but because the thing is really two things. It is an abstraction; from what? From itself. But such an abstraction is inconceivable—unless, possibly, Hegel could have conceived it.

"(2) Let w be the class of all those classes which are not members of themselves. Then, whatever class x may be, 'x is a w' is equivalent to 'x is not an x'. Hence, giving to x the value w, 'w is a w' is equivalent to 'w is not a w'."

The nonsense here involved becomes obvious as soon as we define a class, after the manner of *Principia*, as an extensional function. Then "x is an x" becomes: "if $\phi\hat{x}$ is an extensional function, $\phi\hat{x}$ is part of its extension". This is certainly not the sort of statement we have in mind when we write the abstract form "x is an x". We have in mind an x that denotes the truth-range of a function which may include other extensional functions; but such an included function has another extension, and should be expressed by another symbol.

The only case where "x is an x" is really misleading, is the case of the "class of all classes". It does not seem self-evident that this class cannot be constructed. But the defining function, $\phi(\ )$, would then be an abstraction from *all* defining functions, including $\phi(\ )$; so that again, in one instance it would be an abstraction from itself, which is not acceptable; x may denote *either* the extension of $\phi(\ )$ *or* something constitutive of that extension, but not both; again it is the x, not the ϕ, which is ambiguous. So much for the class of all classes, which is supposed to be self-

[1] This and the following contradictions may be found on pp. 60 and 61 of *Princ. Math.*

including; by analogy, we should of course be unable to construct the function corresponding to a self-excluding class, let alone a class of such classes.

"(3) Let T be the relation which subsists between two relations, R and S, whenever R does not have the relation R to S. Then, whatever relations R and S may be, 'R has the relation T to S' is equivalent to 'R does not have the relation R to S'. Hence, giving the value T to both R and S, 'T has the relation T to T' is equivalent to 'T does not have the relation T to T'."

Now if we regard relations as classes of couples defined by a certain function, we may write: $R = \hat{x}\hat{y}(\phi xy)$, and $S = \hat{u}\hat{v}(\psi uv)$.[1] Thus R is the class of couples, x, y, defined by the function $\phi\hat{x}\hat{y}$. The terms of the relation R are any values of x and y respectively. But if we assert that the relation R holds between R and S, we give to x the value $\hat{x}\hat{y}(\phi xy)$, and assert: "The class of x's and y's of which $\phi(xy)$ is true, and the class of u's and v's of which $\psi(uv)$ is true, together constitute one member of the class of x's and y's of which $\phi(xy)$ is true." Thus we are letting $\hat{x}\hat{y}$ denote (1) the range of truth of $\phi(\hat{x}\hat{y})$ [or to retain our symbolism, $\phi(\ )(\ )$], and—in this case—(2) a part of one of the things in that range. Again, the double meaning is in the argument to one of the functions involved, not in the function itself.

"(4) Burali-Forti's contradiction may be stated as follows: It can be shown that every well-ordered series has an ordinal number, that the series up to and including any given ordinal exceeds the given ordinal by one, and (on certain very natural assumptions) that the series of all ordinals (in order of magnitude) is well-ordered. It follows that the series of all ordinals has an ordinal number, Ω say. But in that case the series of all ordinals including Ω has the ordinal number $\Omega + 1$, which must be greater than Ω. Hence Ω is not the ordinal number of all ordinals."

Now let x = the ordinal number of any well-ordered series, and $(x - n)$, . . . $(x - 1)$, x, $(x + 1)$, . . . $(x + n)$ etc., be the series of ordinals, and Ω the ordinal number of the series of ordinals. Then if we write the abstractive function which defines the class of ordinals, $\phi(\ )$, this is an abstraction from ϕx, ϕy, ϕz, meaning "x is an ordinal number," "y is an ordinal number," etc., and "Ω is an ordinal number" is one instance for the abstraction of $\phi(\ )$. Thus we should have one value for the propositional function "x is an ordinal number" where x denotes an abstraction from a complex wherein it, x, is a constituent; here again the ambiguity of the argument vitiates the proposition we sought to construct.

In like manner we might show that the notation of every illegitimate totality involves a symbol which serves to denote both the extension of some function, and one of its constituent propositions. This is due to the hierarchy of abstractions which is deducible from Mr. Russell's new definition of $\phi\hat{x}$ and his identification of $\phi\hat{x}$ with $\hat{x}(\phi x)$. This interpretation of the hierarchy of types as a hierarchy

[1] *Cf. Princ. Math.*, *21·03.

of abstractions, if correct, serves to establish the type-theory as a *type-theorem*, a necessary proposition in the system of *Principia Mathematica*. How, then, can we account for the fact that the theory appears to rule out certain propositions which common-sense would consider perfectly innocent?

The answer is, that in the original edition of *Principia* we were led to believe that ϕx, ϕy, etc., were *values* for $\phi\hat{x}$, *i.e.*, were ambiguously denoted by it.[1] This would appear to place the two structures, ϕx and $\phi\hat{x}$, upon the same "level of abstraction," with a difference only in their respective degrees of ambiguity, just like "fx" and "fa". Anything, then, that holds for the ambiguous case, holds also for the specific, and in condemning the form $\phi(\phi\hat{x})$ we thought to condemn the form $\phi(\phi x)$. Now in fact there may be paradoxes of the form $\phi(\phi x)$, but they are not due to a confusion of types. The case of the man who says "I am lying," which Mr. Russell quoted as the simplest form of the *Epimenides*, is instructive here. The man who says "I am lying" is not making a statement about *all* his statements, so that his assertion does not involve an illegitimate totality, and is not of the form $\phi(\phi\hat{x})$. He simply asserts "*this* proposition is a lie". Now if "this proposition is a lie" is expressed by ϕx, then x denotes the argument which is: "this proposition". But "this proposition" denotes a certain proposition, namely our original ϕx—"this proposition is a lie". If then we would substitute for x its assigned value, then our proposition, now $\phi(\phi x)$, becomes: "'this proposition is a lie' is a lie," which is no longer—this proposition. "This proposition is a lie" and "'This proposition is a lie' is a lie" are, as a matter of fact, two discrete propositions, and cannot be denoted by the same symbol in the same complex. The fallacy involved is a two-fold interpretation of x, once as x distinct from ϕ, and once as ϕ and x together; ϕx as an interpretation of x would lead to an infinite regress, for in a single structure, if a value is assigned to x, it must be assigned wherever x occurs;[2] consequently we should have to go from ϕx to $\phi(\phi x)$ to $\phi(\phi(\phi x))$ . . . *ad infinitum*. If we assign values to a variable it behoves us to assign them consistently throughout the entire expression. This is a fundamental canon of notation. But the fact that the same x cannot occur as the whole argument and as a constituent in the argument, does not imply that ϕx cannot take, for example, the argument ϕy, with the same ϕ as ϕx. The arguments of the two propositions "This is a lie" and "'This is a lie' is a lie," are different, but the assertion is the same. It is only a misuse of the variable that gives rise to the fallacy here, not the *form* $\phi(\phi x)$. That the exclusion of this form, which is commonly supposed to result from the type-theory, would vitiate many perfectly sane propositions, is obvious. But if our account is true, the theory of types does not properly cover this form.

[1] P. 40.

[2] This is provided in the "axiom of the identification of real variables," (*Princ. Math.*, *1·72).

SUSANNE K. LANGER.