Incomplete symbols: Descriptions

ALFRED NORTH WHITEHEAD
AND BERTRAND RUSSELL

(*1910*)

Whitehead published the first thick volume of *A treatise of universal algebra* in 1898. His pupil Russell published the first thick volume of *The principles of mathematics* in 1903. Then Whitehead was persuaded to abandon plans for his second volume and rather to collaborate with Russell on the second volume of *The principles*. This volume never materialized either. Instead, the collaboration bore as its fruit a decade later a work independent of *The principles*, the three-volume *Principia mathematica*, from which the present selection is drawn. This monumental work is a detailed study of logic and set theory, and a construction, on that basis, of classical mathematics. It came as a fulfillment of what Frege and Peano, in their different ways, had projected and partly carried through.

The present selection from the Introduction to *Principia* is concerned with the logical device of *description*, "$(\imath x)(\varphi x)$", "the object x such that φx". In Peano's notation (*1899*, pp. 22–23) the class whose sole member is x was referred to as ιx, and conversely the sole member x of a class y was $\imath y$; so Whitehead and Russell's "$(\imath x)(\varphi x)$" was an adaptation of Peano's "$\imath x \in \varphi x$", "the member of the class $x \in \varphi x$". Frege had a notation for the same purpose earlier (*1893*, § 11). But there were differences in the handling of the waste cases—cases in which there is not one and only one object x such that φx. Peano left his "$\imath$" unexplained in the waste cases. Frege arbitrarily equated description to class abstraction in the waste cases; "the" object x fulfilling a given condition became for him the class of all such objects x if there were many or none.

Whitehead and Russell contrive to avoid both the incompleteness of Peano's line and the artificiality of Frege's by defining the notation of description for all cases and yet giving it no object in the waste cases. They use contextual definition: they indicate how any sentence containing a description can as a whole be expanded into primitive notation, but they equate the description by itself to no term in primitive notation.

The general idea of contextual definition goes back to Jeremy Bentham (*1843*, vol. 8, pp. 246–247), who called it *paraphrasis* and saw in it a method of accounting for fictions by explaining various purported terms away. A mathematical application of the method became familiar in the differential operators; and it is with this illustration that the selection from Whitehead and Russell begins.

The theory of descriptions that occupies these pages was expounded earlier by Russell (*1905*) in a way that showed motives not unlike Bentham's. Russell was dissatisfied with such theories as "that of Meinong, [which] regards any grammatically correct denoting phrase as

standing for an *object*. Thus 'the present King of France', 'the round square', etc., are supposed to be genuine objects" (*1905*, p. 483). At Whitehead and Russell's hands the method of paraphrasis receives sharper formulation than it had had from Bentham.

A disadvantage of Whitehead and Russell's way of handling descriptions, as compared with Frege's or Peano's, is that it raises the question of *scope*, to which a good part of the appended selection is given over. An advantage, we saw, is completeness without artificiality. Another advantage is that this way of introducing description does not use class abstraction, and hence leaves us free to define the notation "$x \ni \varphi x$" or "$\hat{x}\varphi x$" of class abstraction in terms of description as $(\imath y)(x)(x \in y .\equiv \varphi x)$. But this is an advantage that, in the continuing pages of *Principia mathematica*, Whitehead and Russell do not use. Instead they assume, not quite explicitly, a primitive notation for the abstraction of attributes (or propositional functions), and then with its help they introduce class abstraction by contextual definition (see above, p. 152).

More generally, Whitehead and Russell's way of eliminating descriptions enables us to make a clean sweep not only of class abstracts but of all other sorts of singular terms as well, except variables. Instead of assuming a constant singular term or proper name, say "a", we can always assume or define a predicate "A" that is true only of the object a, and then take "a" itself as short for the eliminable description "$(\imath x)Ax$". Similarly, instead of assuming functors, say the two-place functor "$+$", we can assume or define a three-place predicate "Σ" such that $\Sigma zxy \equiv .\, z = x + y$, and then take "$x + y$" itself as short for the eliminable description "$(\imath z)\Sigma zxy$". An important thing about thus absorbing all singular terms other than variables is that the logic of quantification and identity no longer needs to be conceived otherwise than in its simplest form, involving just predicate letters, variables, quantifiers, truth functions, and "$=$".

W. V. Quine

The text is reproduced here by arrangement with Cambridge University Press.

By an "incomplete" symbol we mean a symbol which is not supposed to have any meaning in isolation, but is only defined in certain contexts. In ordinary mathematics, for example, d/dx and $\int_a^b$ are incomplete symbols: something has to be supplied before we have anything significant. Such symbols have what may be called a "definition in use". Thus, if we put

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \quad \text{Df},$$

we define the *use* of ∇^2, but ∇^2 by itself remains without meaning. This distinguishes such symbols from what (in a generalized sense) we may call *proper names*: "Socrates", for example, stands for a certain man and therefore has a meaning by itself, without the need of any context. If we supply a context, as in "Socrates is mortal", these words express a fact of which Socrates himself is a constituent: there is a certain object, namely Socrates, which does have the property of mortality, and this object is a constituent of the complex fact which we assert when we say "Socrates is mortal". But in other cases this simple analysis fails us. Suppose we say: "The round square does not exist." It seems plain that this is a true proposition, yet we cannot regard it as denying the existence of a certain object called "the round

square". For if there were such an object, it would exist: we cannot first assume that there is a certain object and then proceed to deny that there is such an object. Whenever the grammatical subject of a proposition can be supposed not to exist, without rendering the proposition meaningless, it is plain that the grammatical subject is not a proper name, that is, not a name directly representing some object. Thus, in all such cases, the proposition must be capable of being so analyzed that what was the grammatical subject shall have disappeared. Thus, when we say "the round square does not exist", we may, as a first attempt at such analysis, substitute "it is false that there is an object x which is both round and square". Generally, when "the so-and-so" is said not to exist, we have a proposition of the form[1]

$$\text{“} \sim E!(\imath x)(\varphi x)\text{”},$$

that is,

$$\sim\{(\exists c):\varphi x .\equiv_x. x = c\},$$

or some equivalent. Here the apparent grammatical subject $(\imath x)(\varphi x)$ has completely disappeared; thus, in "$\sim E!(\imath x)(\varphi x)$", $(\imath x)(\varphi x)$ is an *incomplete* symbol.

By an extension of the above argument, it can easily be shown that $(\imath x)(\varphi x)$ is *always* an incomplete symbol. Take, for example, the following proposition: "Scott is the author of Waverley". (Here "the author of Waverley" is $(\imath x)(x$ wrote Waverley).) This proposition expresses an identity; thus, if "the author of Waverley" could be taken as a proper name and supposed to stand for some object c, the proposition would be "Scott is c". But if c is any one except Scott, this proposition is false; while if c *is* Scott, the proposition is "Scott is Scott", which is trivial and plainly different from "Scott is the author of Waverley". Generalizing, we see that the proposition

$$a = (\imath x)(\varphi x)$$

is one which may be true or may be false, but is never merely trivial, like $a = a$; whereas, if $(\imath x)(\varphi x)$ were a proper name, $a = (\imath x)(\varphi x)$ would necessarily be either false or the same as the trivial proposition $a = a$. We may express this by saying that $a = (\imath x)(\varphi x)$ is not a value of the propositional function $a = y$, from which it follows that $(\imath x)(\varphi x)$ is not a value of y. But, since y may be anything, it follows that $(\imath x)(\varphi x)$ is nothing. Hence, since in use it has meaning, it must be an incomplete symbol.

It might be suggested that "Scott is the author of Waverley" asserts that "Scott" and "the author of Waverley" are two names for the same object. But a little reflection will show that this would be a mistake. For if that were the meaning of "Scott is the author of Waverley", what would be required for its truth would be that Scott should have been *called* the author of Waverley: if he had been so called, the proposition would be true, even if some one else had written Waverley; while, if no one called him so, the proposition would be false, even if he had written Waverley. But in fact he was the author of Waverley at a time when no one called him so, and he would not have been the author if every one had called him so but some one else had written Waverley. Thus the proposition "Scott is the author of Waverley" is not a proposition about names, like "Napoleon is Bonaparte"; and this illustrates the sense in which "the author of Waverley" differs from a true proper name.

[1] See *Whitehead and Russell 1910*, pp. 31–32 [[or *1925*, pp. 30–31]].

Thus all phrases (other than propositions) containing the word *the* (in the singular) are incomplete symbols: they have a meaning in use, but not in isolation. For "the author of Waverley" cannot mean the same as "Scott", or "Scott is the author of Waverley" would mean the same as "Scott is Scott", which it plainly does not; nor can "the author of Waverley" mean anything other than "Scott", or "Scott is the author of Waverley" would be false. Hence "the author of Waverley" means nothing.

It follows from the above that we must not attempt to define "$(\imath x)(\varphi x)$", but must define the *uses* of this symbol, that is, the propositions in whose symbolic expression it occurs. Now in seeking to define the uses of this symbol, it is important to observe the import of propositions in which it occurs. Take as an illustration: "The author of Waverley was a poet". This implies (1) that Waverley was written, (2) that it was written by one man, and not in collaboration, (3) that the one man who wrote it was a poet. If any one of these fails, the proposition is false. Thus "the author of 'Slawkenburgius on Noses' was a poet" is false, because no such book was ever written; "the author of 'The Maid's Tragedy' was a poet" is false, because this play was written by Beaumont and Fletcher jointly. These two possibilities of falsehood do not arise if we say "Scott was a poet". Thus our interpretation of the uses of $(\imath x)(\varphi x)$ must be such as to allow for them. Now taking φx to replace "x wrote Waverley", it is plain that any statement apparently about $(\imath x)(\varphi x)$ requires (1) $(\exists x).(\varphi x)$ and (2) $\varphi x.\varphi y .\supset_{x,y}. x = y$; here (1) states that *at least* one object satisfies φx, while (2) states that *at most* one object satisfies φx. The two together are equivalent to

$$(\exists c):\varphi x .\equiv_x. x = c,$$

which we defined as

$$E!(\imath x)(\varphi x).$$

Thus "$E!(\imath x)(\varphi x)$" must be part of what is affirmed by any proposition about $(\imath x)(\varphi x)$. If our proposition is $f\{(\imath x)(\varphi x)\}$, what is further affirmed is fc, if $\varphi x .\equiv_x. x = c$. Thus we have

$$f\{(\imath x)(\varphi x)\} .=: (\exists c):\varphi x .\equiv_x. x = c:fc \quad \text{Df},$$

that is, "the x satisfying φx satisfies fx" is to mean: "There is an object c such that φx is true when, and only when, x is c, and fc is true", or, more exactly: "There is a c such that 'φx' is always equivalent to 'x is c' and fc". In this, "$(\imath x)(\varphi x)$" has completely disappeared; thus "$(\imath x)(\varphi x)$" is merely symbolic, and does not directly represent an object, as single small Latin letters are assumed to do.[2]

The proposition "$a = (\imath x)(\varphi x)$" is easily shown to be equivalent to "$\varphi x .\equiv_x. x = a$". For, by the definition, it is

$$(\exists c):\varphi x .\equiv_x. x = c:a = c,$$

that is, "there is a c for which $\varphi x .\equiv_x. x = c$, and this c is a", which is equivalent to "$\varphi x .\equiv_x. x = a$". Thus "Scott is the author of Waverley" is equivalent to

"'x wrote Waverley' is always equivalent to 'x is Scott'",

that is, "x wrote Waverley" is true when x is Scott and false when x is not Scott.

[2] We shall generally write "$f(\imath x)(\varphi x)$" rather than "$f\{(\imath x)(\varphi x)\}$" in future.

Thus, although "$(\imath x)(\varphi x)$" has no meaning by itself, it may be substituted for y in any propositional function fy, and we get a significant proposition, though not a value of fy.

When $f\{(\imath x)(\varphi x)\}$, as above defined, forms part of some other proposition, we shall say that $(\imath x)(\varphi x)$ has a *secondary* occurrence. When $(\imath x)(\varphi x)$ has a secondary occurrence, a proposition in which it occurs may be true even when $(\imath x)(\varphi x)$ does not exist. This applies, for example, to the proposition: "There is no such person as the King of France". We may interpret this as

$$\sim\{E!(\imath x)(\varphi x)\},$$

or as

$$\sim\{(\exists c).c = (\imath x)(\varphi x)\},$$

if "φx" stands for "x is King of France". In either case, what is asserted is that a proposition p in which $(\imath x)(\varphi x)$ occurs is false, and this proposition p is thus part of a larger proposition. The same applies to such a proposition as the following: "If France were a monarchy, the King of France would be of the House of Orleans".

It should be observed that such a proposition as

$$\sim f\{(\imath x)(\varphi x)\}$$

is ambiguous; it may deny $f\{(\imath x)(\varphi x)\}$, in which case it will be true if $(\imath x)(\varphi x)$ does not exist, or it may mean

$$(\exists c):\varphi x .\equiv_x. x = c :\sim fc,$$

in which case it can be true only if $(\imath x)(\varphi x)$ exists. In ordinary language, the latter interpretation would usually be adopted. For example, the proposition "the King of France is not bald" would usually be rejected as false, being held to mean "the King of France exists and is not bald", rather than "it is false that the King of France exists and is bald". When $(\imath x)(\varphi x)$ exists, the two interpretations of the ambiguity give equivalent results; but, when $(\imath x)(\varphi x)$ does not exist, one interpretation is true and one is false. It is necessary to be able to distinguish these in our notation; and generally, if we have such propositions as

$$\psi(\imath x)(\varphi x) .\supset. p,$$
$$p .\supset. \psi(\imath x)(\varphi x),$$
$$\psi(\imath x)(\varphi x) .\supset. \chi(\imath x)(\varphi x),$$

and so on, we must be able by our notation to distinguish whether the whole or only part of the proposition concerned is to be treated as the "$f(\imath x)(\varphi x)$" of our definition. For this purpose, we will put "$[(\imath x)(\varphi x)]$" followed by dots at the beginning of the part (or whole) which is to be taken as $f(\imath x)(\varphi x)$, the dots being sufficiently numerous to bracket off the $f(\imath x)(\varphi x)$; that is, $f(\imath x)(\varphi x)$ is to be everything following the dots until we reach an equal number of dots not signifying a logical product, or a greater number signifying a logical product, or the end of the sentence, or the end of a bracket enclosing "$[(\imath x)(\varphi x)]$". Thus

$$[(\imath x)(\varphi x)].\psi(\imath x)(\varphi x) .\supset. p$$

will mean

$$(\exists c):\varphi x .\equiv_x. x = c:\psi c :\supset. p,$$

but

$$[(\imath x)(\varphi x)]:\psi(\imath x)(\varphi x) .\supset. p$$

will mean

$$(\exists c):\varphi x .\equiv_x. x = c:\psi c .\supset. p.$$

It is important to distinguish these two, for, if $(\imath x)(\varphi x)$ does not exist, the first is true and the second false. Again

$$[(\imath x)\varphi x)] .\sim\psi(\imath x)(\varphi x)$$

will mean

$$(\exists c):\varphi x .\equiv_x. x = c:\sim\psi c,$$

while

$$\sim\{[(\imath x)(\varphi x)] .\psi(\imath x)(\varphi x)\}$$

will mean

$$\sim\{(\exists c):\varphi x .\equiv_x. x = c:\psi c\}.$$

Here again, when $(\imath x)(\varphi x)$ does not exist, the first is false and the second true.

In order to avoid this ambiguity in propositions containing $(\imath x)(\varphi x)$, we amend our definition, or rather our notation, putting

$$[(\imath x)(\varphi x)]\ f(\imath x)(\varphi x) .=: (\exists c):\varphi x .\equiv_x. x = c:fc \quad \text{Df.}$$

By means of this definition, we avoid any doubt as to the portion of our whole asserted proposition which is to be treated as the "$f(\imath x)(\varphi x)$" of the definition. This portion will be called the *scope* of $(\imath x)(\varphi x)$. Thus in

$$[(\imath x)(\varphi x] .f(\imath x)(\varphi x) .\supset. p$$

the scope of $(\imath x)(\varphi x)$ is $f(\imath x)(\varphi x)$; but in

$$[(\imath x)(\varphi x)]:f(\imath x)(\varphi x) .\supset. p$$

the scope is

$$f(\imath x)(\varphi x) .\supset. p;$$

in

$$\sim\{[(\imath x)(\varphi x)] .f(\imath x)(\varphi x)\}$$

the scope is $f(\imath x)(\varphi x)$; but in

$$[(\imath x)(\varphi x)] .\sim f(\imath x)(\varphi x)$$

the scope is

$$\sim f(\imath x)(\varphi x).$$

It will be seen that, when $(\imath x)(\varphi x)$ has the whole of the proposition concerned for its scope, the proposition concerned cannot be true unless $E!(\imath x)(\varphi x)$; but, when $(\imath x)(\varphi x)$ has only part of the proposition concerned for its scope, it may often be true even when $(\imath x)(\varphi x)$ does not exist. It will be seen further that, when $E!(\imath x)(\varphi x)$, we may enlarge or diminish the scope of $(\imath x)(\varphi x)$ as much as we please without altering the truth value of any proposition in which it occurs.

If a proposition contains two descriptions, say $(\imath x)(\varphi x)$ and $(\imath x)(\psi x)$, we have to distinguish which of them has the larger scope, that is, we have to distinguish

(1) $$[(\imath x)(\varphi x)]:[(\imath x)(\psi x)] .f\{(\imath x)(\varphi x), (\imath x)(\psi x)\},$$

(2) $$[(\imath x)(\psi x)]:[(\imath x)(\varphi x)] .f\{(\imath x)(\varphi x), (\imath x)(\psi x)\}.$$

The first of these, eliminating $(\imath x)(\varphi x)$, becomes

$$(3) \qquad (\exists c):\varphi x .\equiv_x. x = c:[(\imath x)(\psi x)]\ f\{c, (\imath x)(\psi x)\},$$

which, eliminating $(\imath x)(\psi x)$, becomes

$$(4) \qquad (\exists c):. \varphi x .\equiv_x. x = c:.(\exists d):\psi x .\equiv_x. x = c:f(c, d),$$

and the same proposition results if, in (1), we eliminate first $(\imath x)(\psi x)$ and then $(\imath x)(\varphi x)$. Similarly (2) becomes, when $(\imath x)(\varphi x)$ and $(\iota x)(\psi x)$ are eliminated,

$$(5) \qquad (\exists d):.\psi x .\equiv_x. x = d:.(\exists c):\varphi x .=_x. x = c:f(c, d).$$

(4) and (5) are equivalent, so that the truth value of a proposition containing two descriptions is independent of the question which has the larger scope.

It will be found that, in most cases in which descriptions occur, their scope is, in practice, the smallest proposition enclosed in dots or other brackets in which they are contained. Thus, for example,

$$[(\imath x)(\varphi x)].\psi(\imath x)(\varphi x) .\supset. [(\imath x)(\varphi x)].\chi(\imath x)(\varphi x)$$

will occur much more frequently than

$$[(\imath x)(\varphi x)]:\psi(\imath x)(\varphi x) .\supset. \chi(\imath x)(\varphi x).$$

For this reason it is convenient to decide that, when the scope of an occurrence of $(\imath x)(\varphi x)$ is the smallest proposition, enclosed in dots or other brackets, in which the occurrence in question is contained, the scope need not be indicated by "$[(\imath x)(\varphi x)]$". Thus, for example,

$$p .\supset. a = (\imath x)(\varphi x)$$

will mean

$$p .\supset. [(\imath x)(\varphi x)].a = (\imath x)(\varphi x);$$

and

$$p .\supset. (\exists a).a = (\imath x)(\varphi x)$$

will mean

$$p .\supset. (\exists a).[(\imath x)(\varphi x)].a = (\imath x)(\varphi x);$$

and

$$p .\supset. a \neq (\imath x)(\varphi x)$$

will mean

$$p .\supset. [(\imath x)(\varphi x)].\sim\{a = (\imath x)(\varphi x)\};$$

but

$$p .\supset. \sim\{a = (\imath x)(\varphi x)\}$$

will mean

$$p .\supset. \sim\{[(\imath x)(\varphi x)].a = (\imath x)(\varphi x)\}.$$

This convention enables us, in the vast majority of cases that actually occur, to dispense with the explicit indication of the scope of a descriptive symbol; and it will be found that the convention agrees very closely with the tacit conventions of ordinary language on this subject. Thus, for example, if "$(\imath x)(\varphi x)$" is "the so-and-so", "$a \neq (\imath x)(\varphi x)$" is to be read "$a$ is not the so-and-so", which would ordinarily be regarded as implying that "the so-and-so" exists; but "$\sim\{a = (\imath x)(\varphi x)\}$" is to be

read "it is not true that a is the so-and-so", which would generally be allowed to hold if "the so-and-so" does not exist. Ordinary language is, of course, rather loose and fluctuating in its implications on this matter; but, subject to the requirement of definiteness, our convention seems to keep as near to ordinary language as possible.

In the case when the smallest proposition enclosed in dots or other brackets contains two or more descriptions, we shall assume, in the absence of any indication to the contrary, that one which typographically occurs earlier has a larger scope than one which typographically occurs later. Thus

$$(\imath x)(\varphi x) = (\imath x)(\psi x)$$

will mean

$$(\exists c):\varphi x .\equiv_x. x = c:[(\imath x)(\psi x)].c = (\imath x)(\psi x),$$

while

$$(\imath x)(\psi x) = (\imath x)(\varphi x)$$

will mean

$$(\exists d):\psi x .\equiv_x. x = d:[(\imath x)(\varphi x)].(\imath x)(\varphi x) = d.$$

These two propositions are easily shown to be equivalent.