Logic & the Reification of Universals

FROM A LOGICAL POINT OF VIEW

9 Logico-Philosophical Essays

Willard Van Orman Quine
Edgar Pierce Professor of Philosophy
Harvard University

Second Edition, revised

HARVARD UNIVERSITY PRESS
Cambridge, Massachusetts
and London, England

1980

VI

LOGIC AND THE REIFICATION OF UNIVERSALS

1

There are those who feel that our ability to understand general terms, and to see one concrete object as resembling another, would be inexplicable unless there were universals as objects of apprehension. And there are those who fail to detect, in such appeal to a realm of entities over and above the concrete objects in space and time, any explanatory value.

Without settling that issue, it should still be possible to point to certain forms of discourse as *explicitly* presupposing entities of one or another given kind, say universals, and purporting to treat of them; and it should be possible to point to other forms of discourse as not explicitly presupposing those entities. Some criterion to this purpose, some standard of ontological commitment, is needed if we are ever to say meaningfully that a given theory depends on or dispenses with the assumption of such and such objects. Now we saw earlier[1] that such a criterion is to be found not in the singular terms of the given discourse, not in the purported names, but rather in quantification. We shall be occupied in these pages with a closer examination of the point.

The quantifiers '$(\exists x)$' and '(x)' mean 'there is some entity x such that' and 'each entity x is such that'. The letter 'x' here, called a bound variable, is rather like a pronoun; it is used in

[1] Pp. 12ff.

the quantifier to key the quantifier for subsequent cross-reference, and then it is used in the ensuing text to refer back to the appropriate quantifier. The connection between quantification and entities outside language, be they universals or particulars, consists in the fact that the truth or falsity of a quantified statement ordinarily depends in part on what we reckon into the range of entities appealed to by the phrases 'some entity x' and 'each entity x'—the so-called range of values of the variable. That classical mathematics treats of universals, or affirms that there are universals, means simply that classical mathematics requires universals as values of its bound variables. When we say, for example,

$$(\exists x)(x \text{ is prime} . x > 1{,}000{,}000),$$

we are saying that *there is* something which is prime and exceeds a million; and any such entity is a number, hence a universal. In general, *entities of a given sort are assumed by a theory if and only if some of them must be counted among the values of the variables in order that the statements affirmed in the theory be true.*

I am not suggesting a dependence of being upon language. What is under consideration is not the ontological state of affairs, but the ontological commitments of a discourse. What there is does not in general depend on one's use of language, but what one says there is does.

The above criterion of ontological commitment applies in the first instance to discourse and not to men. One way in which a man may fail to share the ontological commitments of his discourse is, obviously, by taking an attitude of frivolity. The parent who tells the Cinderella story is no more committed to admitting a fairy godmother and a pumpkin coach into his own ontology than to admitting the story as true. Another and more serious case in which a man frees himself from ontological commitments of his discourse is this: he shows how some particular use which he makes of quantification, involving a prima facie commitment to certain objects, can be expanded into an idiom innocent of such commitments. (See, for example, §4, below.) In this event the seemingly presupposed objects may justly be said to have been explained away as convenient fictions, manners of speaking.

Contexts of quantification, '$(x)(\ldots x \ldots)$' and '$(\exists x)(\ldots x \ldots)$', do not exhaust the ways in which a variable 'x' may turn up in discourse. The variable is also essential to the idiom of singular description 'the object x such that . . .', the idiom of class abstraction 'the class of all objects x such that . . .', and others. However, the quantificational use of variables is exhaustive in the sense that all use of bound variables is *reducible* to this sort of use. Every statement containing a variable can be translated, by known rules, into a statement in which the variable has only the quantificational use.[2] All other uses of bound variables can be explained as abbreviations of contexts in which the variables figure solely as variables of quantification.

It is equally true that any statement containing variables can be translated, by other rules, into a statement in which variables are used solely for class abstraction;[3] and, by still other rules, into a statement in which variables are used solely for functional abstraction (as in Church [1]). Whichever of these roles of variables be taken as fundamental, we can still hold to the criterion of ontological commitment italicized above.

An ingenious method invented by Schönfinkel, and developed by Curry and others, gets rid of variables altogether by recourse to a system of constants, called combinators, which express certain logical functions. The above criterion of ontological commitment is of course inapplicable to discourse constructed by means of combinators. Once we know the systematic method of translating back and forth between statements which use combinators and statements which use variables, however, there is no difficulty in devising an equivalent criterion of ontological commitment for combinatory discourse. The entities presupposed by statements which use combinators turn out, under such reasoning, to be just the entities that must be reckoned as arguments or values of functions in order that the statements in question be true.

[2] See above, pp. 85ff.
[3] See above, pp. 94f.

But it is to the familiar quantificational form of discourse that our criterion of ontological commitment primarily and fundamentally applies. To insist on the correctness of the criterion in this application is, indeed, merely to say that no distinction is being drawn between the 'there are' of 'there are universals', 'there are unicorns', 'there are hippopotami', and the 'there are' of '$(\exists x)$', 'there are entities x such that'. To contest the criterion, as applied to the familiar quantificational form of discourse, is simply to say either that the familiar quantificational notation is being re-used in some new sense (in which case we need not concern ourselves) or else that the familiar 'there are' of 'there are universals' et al. is being re-used in some new sense (in which case again we need not concern ourselves).

If what we want is a standard for our own guidance in appraising the ontological commitments of one or another of our theories, and in altering those commitments by revision of our theories, then the criterion at hand well suits our purposes; for the quantificational form is a convenient standard form in which to couch any theory. If we prefer another language form, for example, that of combinators, we can still bring our criterion of ontological commitment to bear in so far as we are content to accept appropriate systematic correlations between idioms of the aberrant language and the familiar language of quantification.

Polemical use of the criterion is a different matter. Thus, consider the man who professes to repudiate universals but still uses without scruple any and all of the discursive apparatus which the most unrestrained of platonists might allow himself. He may, if we train our criterion of ontological commitment upon him, protest that the unwelcome commitments which we impute to him depend on unintended interpretations of his statements. Legalistically his position is unassailable, as long as he is content to deprive us of a translation without which we cannot hope to understand what he is driving at. It is scarcely cause for wonder that we should be at a loss to say what objects a given discourse presupposes that there are, failing all notion of how to translate that discourse into the sort of language to which 'there is' belongs.

Also there are the philosophical champions of ordinary language. Their language is emphatically one to which 'there is' belongs, but they look askance at a criterion of ontological commitment which turns on a real or imagined translation of statements into quantificational form. The trouble this time is that the idiomatic use of 'there is' in ordinary language knows no bounds comparable to those that might reasonably be adhered to in scientific discourse painstakingly formulated in quantificational terms. Now a philological preoccupation with the unphilosophical use of words is exactly what is wanted for many valuable investigations, but it passes over, as irrelevant, one important aspect of philosophical analysis—the creative aspect, which is involved in the progressive refinement of scientific language. In this aspect of philosophical analysis any revision of notational forms and usages which will simplify theory, any which will facilitate computations, any which will eliminate a philosophical perplexity, is freely adopted as long as all statements of science can be translated into the revised idiom without loss of content germane to the scientific enterprise. Ordinary language remains indeed fundamental, not only genetically but also as a medium for the ultimate clarification, by however elaborate paraphrase, of such more artificial usages. But it is not with ordinary language, it is rather with one or another present or proposed refinement of scientific language, that we are concerned when we expound the laws of logical inference or such analyses as Frege's of the integer, Dedekind's of the real number, Weierstrass's of the limit, or Russell's of the singular description.[4] And it is only in this spirit, in reference to one or another real or imagined logical schematization of one or another part or all of science, that we can with full propriety inquire into ontological presuppositions. The philosophical devotees of ordinary language are right in doubting the final adequacy of any criterion of the ontological presuppositions of ordinary language, but they are wrong in supposing that there is no more to be said on the philosophical question of ontological presuppositions.

[4] See below, pp. 165ff.

In a loose way we often can speak of ontological presuppositions at the level of ordinary language, but this makes sense just in so far as we have in mind some likeliest, most obvious way of schematizing the discourse in question along quantificational lines. It is here that the 'there is' of ordinary English lends its services as a fallible guide—an all too fallible one if we pursue it purely as philologists, unmindful of the readiest routes of logical schematization.

Relative to a really alien language L it may happen, despite the most sympathetic effort, that we cannot make even the roughest and remotest sense of ontological commitment. There may well be no objective way of so correlating L with our familiar type of language as to determine in L any firm analogue of quantification, or 'there is'. Such a correlation might be out of the question even for a man who has a native fluency in both languages and can interpret back and forth in paragraph units at a business level. In this event, to seek the ontological commitments of L is simply to project a provincial trait of the conceptual scheme of our culture circle beyond its range of significance. Entity, objectuality, is foreign to the L-speaker's conceptual scheme.

2

In the logic of quantification, as it is ordinarily set up, principles are propounded in this style:

(1) $$[(x)(Fx \supset Gx) \,.\, (\exists x)Fx] \supset (\exists x)Gx.$$

'Fx' and 'Gx' stand in place of any sentences, for example, 'x is a whale' and 'x swims'. The letters 'F' and 'G' are sometimes viewed as variables taking attributes or classes as values, for example, whalehood and swimmingness, or whalekind and the class of swimming things. Now what sets attributes apart from classes is merely that whereas classes are identical when they have the same members, attributes may be distinct even though present in all and only the same things. Consequently, if we apply the maxim of identification of indiscernibles[5] to quantifica-

[5] See above, p. 71.

tion theory, we are directed to construe classes rather than attributes as the values of '*F*', '*G*', etc. The constant expressions which '*F*', '*G*', etc. stand in place of, then, namely, predicates or general terms such as 'is a whale' and 'swims', come to be regarded as names of classes; for the things in place of whose names variables stand are values of the variables. To Church [6] is due the interesting further suggestion that whereas predicates name classes, they may be viewed as having attributes rather as their meanings.

But the best course is yet another. We can look upon (1) and similar valid forms simply as schemata or diagrams embodying the form of each of various true statements, for example:

$$(2)\quad [(x)(x \text{ has mass} \supset x \text{ is extended}) \,.\, (\exists x)(x \text{ has mass})] \supset (\exists x)(x \text{ is extended}).$$

There is no need to view the 'has mass' and 'is extended' of (2) as names of classes or of anything else, and there is no need to view the '*F*' and '*G*' of (1) as variables taking classes or anything else as values. For let us recall our criterion of ontological commitment: an entity is presupposed by a theory if and only if it is needed among the values of the bound variables in order to make the statements affirmed in the theory true. '*F*' and '*G*' are not bindable variables, and hence need be regarded as no more than dummy predicates, blanks in a sentence diagram.

In the most elementary part of logic, namely, the logic of truth functions,[6] principles are commonly propounded with '*p*', '*q*', etc. taking the place of component statements; for example, '$[(p \supset q) \,.\, \sim q] \supset \sim p$'. The letters '*p*', '*q*', etc. are sometimes viewed as taking entities of some sort as values; and, since the constant expressions which '*p*', '*q*', etc. stand in place of are statements, those supposed values must be entities whereof statements are names. These entities have sometimes been called *propositions*. In this usage the word 'proposition' is not a synonym of 'statement' (as it commonly is), but refers rather to hypothetical abstract entities of some sort. Alternatively,

[6] See above, p. 84.

notably by Frege [3], statements have been taken to name always just one or the other of two entities, the so-called truth values: the true and the false. Both courses are artificial, but, of the two, Frege's is preferable for its conformity to the maxim of the identification of indiscernibles. Propositions, if one must have them, are better viewed as *meanings* of statements, as Frege pointed out, not as what are named by statements.

But the best course is to revert to the common-sense view, according to which names are one sort of expression and statements another. There is no need to view statements as names, nor to view '*p*', '*q*', etc. as variables which take entities named by statements as values; for '*p*', '*q*', etc. are not used as bound variables subject to quantifiers. We can view '*p*', '*q*', etc. as schematic letters comparable to '*F*', '*G*', etc.; and we can view '$[(p \supset q) \,.\, \sim q] \supset \sim p$', like (1), not as a sentence but as a schema or diagram such that all actual statements of the depicted form are true. The schematic letters '*p*', '*q*', etc. stand in schemata to take the place of component statements, just as the schematic letters '*F*', '*G*', etc. stand in schemata to take the place of predicates; and there is nothing in the logic of truth functions or quantification to cause us to view statements or predicates as names of any entities, or to cause us to view these schematic letters as variables taking any such entities as values. It is only the bound variable that demands values.

Let us interrupt our progress long enough to become quite clear on essential distinctions. Consider the expressions:

$$x + 3 > 7, \qquad (x)(Fx \supset p).$$

The former of these is a sentence. It is not indeed a *closed* sentence, or statement, because of the free 'x'; but it is an open sentence, capable of occurring within a context of quantification to form part of a statement. The other expression, '$(x)(Fx \supset p)$', is not a sentence at all, but a schema, if the attitude is adopted toward '*F*' and '*p*' which was recommended in the preceding paragraph. The schema '$(x)(Fx \supset p)$' cannot be imbedded within quantification to form part of a statement, for schematic letters are not bindable variables.

The letter 'x' is a bindable variable—one whose values, we may temporarily suppose for purposes of the example '$x + 3 > 7$', are numbers. The variable stands in *place* of *names* of numbers, for example, Arabic numerals; the *values* of the variable are the numbers themselves. Now just as the letter 'x' stands in place of numerals (and other names of numbers), so the letter 'p' stands in place of statements (and sentences generally). If statements, like numerals, were thought of as names of certain entities, and 'p', like 'x', were thought of as a bindable variable, then the *values* of 'p' would be such entities as statements were names of. But if we treat 'p' as a schematic letter, an unbindable dummy statement, then we drop the thought of namehood of statements. It remains true that 'p' stands in place of statements as 'x' stands in place of numerals; but whereas the bindable 'x' has numbers as values, the unbindable 'p' does not have values at all. Letters qualify as genuine variables, demanding a realm of objects as their values, only if it is permissible to bind them so as to produce actual statements about such objects.

'F' is on a par with 'p'. If predicates are thought of as names of certain entities and 'F' is treated as a bindable variable, then the values of 'F' are such entities as predicates are names of. But if we treat 'F' as a schematic letter, an unbindable dummy predicate, then we drop the thought of namehood of predicates, and of values for 'F'. 'F' simply stands in place of predicates; or, to speak in more fundamental terms, 'Fx' stands in place of sentences.

If we did not care eventually to use 'x' explicitly or implicitly in quantifiers, then the schematic status urged for 'p' and 'F' would be equally suited to 'x'. This would mean treating 'x' in '$x + 3 > 7$' and similar contexts as a dummy numeral but dropping the thought of there being numbers for numerals to name. In this event '$x + 3 > 7$' would become, like '$(x)(Fx \supset p)$', a mere schema or dummy statement, sharing the form of genuine statements (such as '$2 + 3 > 7$') but incapable of being quantified into a statement.

Both of the foregoing expressions '$x + 3 > 7$' and '$(x)(Fx \supset p)$' are radically different in status from such expres-

sions as:

(3) $$(\exists \alpha)(\phi \vee \psi)$$

in the sense of Essay V. (3) occupies, so to speak, a semantical level next above that of '$x + 3 > 7$' and '$(x)(Fx \supset p)$': it stands as a name *of* a sentence, or comes to do so as soon as we specify a particular choice of expressions for the Greek letters to refer to. A schema such as '$(x)(Fx \supset p)$', on the contrary, is not a name of a sentence, not a name of anything; it is *itself* a pseudo-sentence, designed expressly to manifest a form which various sentences manifest. Schemata are to sentences not as names to their objects, but as slugs to nickels.

The Greek letters are, like 'x', variables, but variables within a portion of language specially designed for talking *about* language. We lately thought about 'x' as a variable which takes numbers as values, and thus stands in place of names of numbers; now correspondingly the Greek letters are variables which take sentences or other expressions as values, and thus stand in place of *names* (for example, quotations) *of* such expressions. Note that the Greek letters are genuine bindable variables, accessible to such verbally phrased quantifiers as 'no matter what statement ϕ may be', 'there is a statement ψ such that'.

Thus 'ϕ' contrasts with 'p' in two basic ways. First, 'ϕ' is a variable, taking sentences as values; 'p', construed schematically, is not a variable (in the value-taking sense) at all. Second, 'ϕ' is grammatically substantival, occupying the place of names of sentences; 'p' is grammatically sentential, occupying the place of sentences.

This latter contrast is dangerously obscured by the usage (3), which shows the Greek letters 'ϕ' and 'ψ' in sentential rather than substantival positions. But this usage would be nonsense except for the special and artificial convention of Essay V (p. 83) concerning the imbedding of Greek letters among signs of the logical language. According to that convention, (3) is shorthand for the unmisleading substantive:

> the result of putting the variable α and the sentences ϕ and ψ in the respective blanks of '$(\exists\ \ )(\ \ \ \vee\ \ \ )$'.

Here the Greek letters clearly occur in noun positions (referring *to* a variable and to two statements), and the whole is a noun in turn. In some of my writings, for example [1], I have insisted on fitting the misleading usage (3) with a safety device in the form of a modified type of quotation marks, thus:

$$\ulcorner(\exists\alpha)(\phi \vee \psi)\urcorner.$$

These marks rightly suggest that the whole is, like an ordinary quotation, a substantive which refers *to* an expression; also they conspicuously isolate those portions of text in which the combined use of Greek letters and logical signs is to be oddly construed. In most of the literature, however, these quasi-quotation marks are omitted. The usage of most logicians who take care to preserve the semantic distinctions at all is that exemplified by Essay V (though commonly with German or boldface Latin letters instead of Greek).

So much for the usage of Greek letters. It will recur as a practical expedient in §§5-6, but its present relevance is simply its present irrelevance. The distinction which properly concerns us in the present pages, that between sentence and schema, is not a distinction between the use and mention of expressions; its significance lies elsewhere altogether. The significance of preserving a schematic status for '*p*', '*q*', etc. and '*F*', '*G*', etc., rather than treating those letters as bindable variables, is that we are thereby (a) forbidden to subject those letters to quantification, and (b) spared viewing statements and predicates as names of anything.

3

The reader must surely think by now that the recommendation of a schematic status for '*p*', '*q*', etc. and '*F*', '*G*', etc. is prompted purely by a refusal to admit entities such as classes and truth values. But this is not true. There can be good cause, as we shall see presently, for admitting such entities, and for admitting names of them, and for admitting bindable variables which take such entities—classes, anyway—as values. My present objection is only against treating statements and predicates themselves as names of such or any entities, and thus identifying the '*p*', '*q*', etc. of truth-function theory and the '*F*', '*G*', etc. of quantification theory with bindable variables. For bindable variables we have '*x*', '*y*', etc., and, if a distinction is wanted between variables for individuals and variables for classes or truth values, we can add distinctive alphabets; but there are reasons for preserving a schematic status for '*p*', '*q*', etc. and '*F*', '*G*', etc.

One reason is that to construe '*Fx*' as affirming membership of *x* in a class can, in many theories of classes, lead to a technical impasse. For there are theories of classes in which not every expressible condition on *x* determines a class, and theories in which not every object is eligible for membership in classes.[7] In such a theory '*Fx*' can represent any condition whatever on any object *x*, whereas '$x \in y$' cannot.

But the main disadvantage of assimilating schematic letters to bound variables is that it leads to a false accounting of the ontological commitments of most of our discourse. When we say that some dogs are white,

(4) $(\exists x)(x \text{ is a dog} \,.\, x \text{ is white})$,

we do not commit ourselves to such abstract entities as dogkind or the class of white things.[8] Hence it is misleading to construe the words 'dog' and 'white' as names of such entities. But we do just that if in representing the form of (4) as '$(\exists x)(Fx \,.\, Gx)$' we think of '*F*' and '*G*' as bindable class variables.

We can of course switch to the explicit form '$(\exists x)(x \in y \,.\, x \in z)$' whenever we really want class variables available for binding. (Also we may use, instead of '*y*' and '*z*', a distinctive style of variables for classes.) Though we do not recognize the general terms 'dog' and 'white' as names of dogkind and the class of white things, genuine names of those abstract entities are not far to seek, namely, the singular terms 'dogkind' and 'the class of white things'. Singular terms naming entities are

[7] See, for example, pp. 92, 96ff above.
[8] See above, p. 13.

quite properly substituted for variables which admit those entities as values; and accordingly we have:

(5) $(\exists x)(x \in \text{dogkind} \,.\, x \in \text{class of white things})$

as an instance of the form '$(\exists x)(x \in y \,.\, x \in z)$'. (5) is also, like (4), an instance of the form '$(\exists x)(Fx \,.\, Gx)$'; but (4) is not an instance of the form '$(\exists x)(x \in y \,.\, x \in z)$'.

I grant that (4) and (5) as wholes are equivalent statements. But they differ in that (4) belongs squarely to the part of language which is neutral on the question of class existence, whereas (5) is tailored especially to fit that higher part of language in which classes are assumed as values of variables. (5) itself just happens to be a degenerate specimen of that higher part of language, in two respects; it actually contains no quantification over classes, and taken as a whole statement it is equivalent to (4).

The assimilation of schematic letters to bound variables, against which I have been inveighing, must indeed be conceded some utility if we want to slip from the ontologically innocent domain of elementary logic into a theory of classes or other abstract entities with a minimum of notice. This could be found desirable either from an unworthy motive of concealment or from a worthier motive of speculating on origins. Acting from the latter motive, I shall in fact exploit the procedure in §§4–5. But the procedure is useful for this purpose precisely because of its faults.

The fact that classes *are* universals, or abstract entities, is sometimes obscured by speaking of classes as mere aggregates or collections, thus likening a class of stones, say, to a heap of stones. The heap is indeed a concrete object, as concrete as the stones that make it up; but the class of stones in the heap cannot properly be identified with the heap. For, if it could, then by the same token another class could be identified with the same heap, namely, the class of molecules of stones in the heap. But actually these classes have to be kept distinct; for we want to say that the one has just, say, a hundred members, while the other has trillions. Classes, therefore, are abstract entities; we may call them aggregates or collections if we like, but they are universals. That is, if there *are* classes.

There are occasions which call quite directly for discourse about classes.[9] One such occasion arises when we define ancestor in terms of parent, by Frege's method: x is ancestor of y if x belongs to *every class* which contains y and all parents of its own members.[10] There is thus serious motive for quantification over classes; and, to an equal degree, there is a place for singular terms which name classes—such singular terms as 'dogkind' and 'the class of Napoleon's ancestors'.

To withhold from general terms or predicates the status of names of classes is not to deny that there are often (or always, apart from the class-theoretic universes noted two pages back) certain classes connected with predicates otherwise than in the fashion of being named. Occasions arise for speaking of the *extension* of a general term or predicate—the class of all things of which the predicate is true. One such occasion arises when we treat the topic of validity of schemata of pure quantification theory; for a quantificational schema is valid when it comes out true for all values of its free (but bindable) variables under all assignments of classes as extensions of the schematic predicate letters. The general theory of quantificational validity thus appeals to classes, but the individual statements represented by the schemata of quantification theory need not; the statement (4) involves, of itself, no appeal to the abstract extension of a predicate.

Similarly there is occasion in the theory of validity to speak of truth values of statements, for example, in defining truth-functional validity. But there is no need to treat statements as names of these values, nor as names at all. When we simply affirm a statement we do not thereby appeal to any such entity as a truth value, unless the statement happens to have that special subject matter.

It can indeed prove convenient and elegant in special systems to reconstrue statements as names—for example, of 2 and 1, as

[9] See above, pp. 12ff.
[10] Note the analogy between this definition and (3) of p. 98.

in Church's system [1]. This is perhaps better regarded as a matter of making names of 2 and 1 serve the purpose of statements, for the special system; and I have no quarrel with it. Similarly Frege may be represented as making his singular terms, plus membership, do the work of general terms; and with this again, as a means merely of absorbing lower logic into a particular system of higher logic for the sake of elegance, there is no quarrel. Special systems aside, however, it is obviously desirable to analyze discourse in such a way as not to impute special ontological presuppositions to portions of discourse which are innocent of them.

The bulk of logical reasoning takes place on a level which does not presuppose abstract entities. Such reasoning proceeds mostly by quantification theory, the laws of which can be represented through schemata involving no quantification over class variables. Much of what is commonly formulated in terms of classes, relations, and even number, can be easily reformulated schematically within quantification theory plus perhaps identity theory.[11] Thus I consider it a defect in an all-purpose formulation of the theory of reference if it represents us as referring to abstract entities from the very beginning rather than only where there is a real purpose in such reference. Hence my wish to keep general terms distinct from abstract singular terms.

Even in the theory of validity it happens that the appeal to truth values of statements and extensions of predicates can finally be eliminated. For truth-functional validity can be redefined by the familiar tabular method of computation, and validity in quantification theory can be redefined simply by appeal to the rules of proof (since Gödel [1] has proved them complete). Here is a good example of the elimination of ontological presuppositions, in one particular domain.

In general it is important, I think, to show how the purposes of a certain segment of mathematics can be met with a reduced ontology, just as it is important to show how an erstwhile nonconstructive proof in mathematics can be accomplished by constructive means. The interest in progress of this type is no more

[11] See below, p. 128.

dependent upon an out-and-out intolerance of abstract entities than it is upon an out-and-out intolerance of nonconstructive proof. The important thing is to understand our instrument; to keep tab on the diverse presuppositions of diverse portions of our theory, and reduce them where we can. It is thus that we shall best be prepared to discover, eventually, the over-all dispensability of some assumption that has always rankled as *ad hoc* and unintuitive.

4

It may happen that a theory dealing with nothing but concrete individuals can conveniently be reconstrued as treating of universals, by the method of identifying indiscernibles. Thus, consider a theory of bodies compared in point of length. The values of the bound variables are physical objects, and the only predicate is 'L', where 'Lxy' means 'x is longer than y'. Now where $\sim Lxy \,.\, \sim Lyx$, anything that can be truly said of x within this theory holds equally for y and vice versa. Hence it is convenient to treat '$\sim Lxy \,.\, \sim Lyx$' as '$x = y$'. Such identification amounts to reconstruing the values of our variables as universals, namely, lengths, instead of physical objects.

Another example of such identification of indiscernibles is obtainable in the theory of *inscriptions*, a formal syntax in which the values of the bound variables are concrete inscriptions The important predicate here is 'C', where '$Cxyz$' means that x consists of a part notationally like y followed by a part notationally like z. The condition of interchangeability or indiscernibility in this theory proves to be notational likeness, expressible thus:

$$(z)(w)(Cxzw \equiv Cyzw \,.\, Czxw \equiv Czyw \,.\, Czwx \equiv Czwy).$$

By treating this condition as '$x = y$', we convert our theory of inscriptions into a theory of notational forms, where the values of the variables are no longer individual inscriptions, but the abstract notational shapes of inscriptions.

This method of abstracting universals is quite reconcilable with nominalism, the philosophy according to which there are really no universals at all. For the universals may be regarded as entering here merely as a manner of speaking—through the metaphorical use of the identity sign for what is really not identity but sameness of length, in the one example, or notational likeness in the other example. In abstracting universals by identification of indiscernibles, we do no more than rephrase the same old system of particulars.

Unfortunately, though, this innocent kind of abstraction is inadequate to abstracting any but mutually exclusive classes. For when a class is abstracted by this method, what holds it together is the indistinguishability of its members by the terms of the theory in question; so any overlapping of two such classes would fuse them irretrievably into a single class.

Another and bolder way of abstracting universals is by admitting into quantifiers, as bound variables, letters which had hitherto been merely schematic letters involving no ontological commitments. Thus if we extend truth-function theory by introducing quantifiers '(p)', '(q)', '$(\exists p)$', etc., we can then no longer dismiss the statement letters as schematic. Instead we must view them as variables taking appropriate entities as values, namely, propositions or, better, truth values, as is evident from the early pages of this essay. We come out with a theory involving universals, or anyway abstract entities.

Actually, though, even the quantifiers '(p)' and '$(\exists p)$' happen to be reconcilable with nominalism if we are working in an extensional system.[12] For, following Tarski [2], we can construe '$(p)(\ldots p\ldots)$' and '$(\exists p)(\ldots p\ldots)$' (where '$\ldots p\ldots$' is any context containing 'p' in the position of a component statement) as the conjunction and alternation of '$\ldots S\ldots$' and '$\ldots \sim S\ldots$', where 'S' is short for some specific statement arbitrarily chosen. If we are working in an extensional system, it can be proved that this artificial way of defining quantification of 'p', 'q', etc. fulfills all the appropriate laws. What seemed to be quantified discourse about propositions or truth values is thereby legitimized, from a nominalist point of view, as a figure of speech.

[12] On extensionality see above, p. 30. For a discussion of nonextensional systems see Essay VIII.

What seemed to be discourse in which statements figured as names is explained as a picturesque transcription of discourse in which they do not.

But abstraction by binding schematic letters is not always thus easily reconcilable with nominalism. If we bind the schematic letters of quantification theory, we achieve a reification of universals which no device analogous to Tarski's is adequate to explaining away. These universals are entities whereof predicates may thenceforward be regarded as names. They might, as noted in §2, be taken as attributes or as classes, but better as classes.

In §3 strong reasons were urged for maintaining a notational distinction between schematic predicate letters, such as the 'F' of 'Fx', and bindable variables used in connection with 'ϵ' to take classes as values. The reasons were reasons of logical and philosophical clarity. Now for those very same reasons, seen in reverse, it can be suggestive to rub out the distinction if we are interested in the genetic side. The ontologically crucial step of positing a universe of classes or other abstract entities can be made to seem a small step, rather naturally taken, if represented as a mere matter of letting erstwhile schematic letters creep into quantifiers. Thus it was that 'p' was admitted unchanged into quantifiers a few paragraphs back. Similarly, in the spirit of an imaginative reënactment of the genesis of class theory, let us now consider in detail how class theory proceeds from quantification theory by binding erstwhile schematic predicate letters.

5

First we must get a closer view of quantification theory. Quantificational schemata are built up of schematic components 'p', 'q', 'Fx', 'Gx', 'Gy', 'Fxy', etc. by means of quantifiers '(x)', '(y)', '$(\exists x)$', etc. and the truth-functional operators '$\sim$', '$.$', '$\vee$', '$\supset$', '$\equiv$'.[13] Various systematizations of quantification theory are known which are complete, in the sense that all the valid schemata are theorems. (See above, §3). One such system is

[13] See above, pp. 83f.

constituted by the rules R1, R2, R4, and R5, of Essay V, above, if we reconstrue the 'ϕ', 'ψ', 'χ', and 'ω' thereof as referring to quantificational schemata. The definitions D1-6 of that essay must be included.

A conspicuous principle of quantification theory is that for all occurrences of a predicate letter followed by variables we can substitute any one condition on those variables. For 'Fx' we can substitute any schema, for example, '$(y)(Gx \supset Hyx)$', provided that for 'Fz', 'Fw', etc. we make parallel substitutions '$(y)(Gz \supset Hyz)$', '$(y)(Gw \supset Hyw)$', etc.[14] This principle of substitution has not had to be assumed along with R1, R2, R4, and R5, simply because its use can in theory always be circumvented as follows: instead, for example, of substituting '$(y)(Gx \supset Hyx)$' for 'Fx' in a theorem ϕ to get a theorem ψ, we can always get ψ by repeating the proof of ϕ itself with '$(y)(Gx \supset Hyx)$' in place of 'Fx'.

Another conspicuous principle of quantification theory is that of *existential generalization*, which carries us from a theorem ϕ to a theorem $(\exists x)\psi$ where ϕ is like ψ except for containing free occurrences of 'y' in all the positions in which ψ contains free occurrences of 'x'. For example, from '$Fy \equiv Fy$' existential generalization yields '$(\exists x)(Fy \equiv Fx)$'. Now this principle has not had to be assumed along with R1, R2, R4, and R5, simply because whatever can be done by use of it can be done also by a devious series of applications of R1, R2, and R4 (and D1-6).

There is no need to favor R1, R2, R4, and R5 as the basic principles for generating valid quantificational schemata. They happen to be an adequate set of rules, but there are also alternative choices that would be adequate;[15] some such choices include substitution or existential generalization as basic, to the exclusion of one or another of R1, R2, R4, and R5.

Now the maneuver of extending quantification to predicate letters, as a means of expanding quantification theory into class theory, can be represented as a provision merely to allow predicate letters all privileges of the variables 'x', 'y', etc. Let us see how this provision works. To begin with, the quantificational schema '$(y)(Gy \equiv Gy)$' is obviously valid and hence must be obtainable as a theorem of pure quantification theory. Now our new provision for granting 'F' and 'G' the privileges of ordinary variables allows us to apply existential generalization to '$(y)(Gy \equiv Gy)$' in such fashion as to obtain '$(\exists F)(y)(Fy \equiv Gy)$'. From this in turn, by substitution, we get $(\exists F)(y)(Fy \equiv \phi)$ where ϕ is any desired condition on y.

'F', admitted thus into quantifiers, acquires the status of a variable taking classes as values; and the notation 'Fy' comes to mean that y is a member of the class F. So the above result $(\exists F)(y)(Fy \equiv \phi)$ is recognizable as R3 of Essay V.[16]

Such extension of quantification theory, simply by granting the predicate variables all privileges of 'x', 'y', etc., would seem a very natural way of proclaiming a realm of universals mirroring the predicates or conditions that can be written in the language. Actually, however, it turns out to proclaim a realm of classes *far wider* than the conditions that can be written in the language. This result is perhaps unwelcome, for surely the intuitive idea underlying the positing of a realm of universals is merely that of positing a reality behind linguistic forms. The result is, however, forthcoming; we can obtain it as a corollary of the theorem of Cantor mentioned earlier.[17] Cantor's proof can be carried out within the extension of quantification theory under consideration, and from his theorem it follows that there must be classes, in particular classes of linguistic forms, having no linguistic forms corresponding to them.

But this is nothing to what *can* be shown in the theory under consideration. For we have seen that the theory is adequate to R1-5, including R3; and we saw in Essay V that R1-5 lead to Russell's paradox.

Classical mathematics has roughly the above theory as its

[14] For a more rigorous formulation of this rule see my [2], §25.

[15] For example, see Hilbert and Ackermann, ch. 3, §5; Quine [1], p. 88; [2], pp. 157-161, 191.

[16] See p. 89 above. The hypothesis of R3, namely, that ϕ lack 'x' (or now 'F'), is strictly needed because of restrictions which enter into any rigorous formulation of the rule of substitution whereby ϕ was just now substituted for 'Gy'.

[17] P. 92n.

foundation, subject, however, to one or another arbitrary restriction, of such kind as to restore consistency without disturbing Cantor's result. Various such restrictions were reviewed earlier.[18] Incidentally, the notation just now developed can be cut down by dropping the polyadic use of bindable predicate variables (such as '*F*' in '*Fxy*'), since relations are constructible as in Essay V from classes; and the residual forms '*Fx*', '*Fy*', '*Gx*', etc., with bindable '*F*', '*G*', etc., can be rewritten as '$x \epsilon z$', '$y \epsilon z$', '$x \epsilon w$', etc. in conformity with what was urged early in the present essay. We come out with the notation of Essay V. But in any case universals are irreducibly presupposed. The universals posited by binding the predicate letters have never been explained away in terms of any mere convention of notational abbreviation, such as we were able to appeal to in earlier less sweeping instances of abstraction.

The classes thus posited are, indeed, all the universals that mathematics needs. Numbers, as Frege showed, are definable as certain classes of classes. Relations, as noted, are likewise definable as certain classes of classes. And functions, as Peano emphasized, are relations. Classes are enough to worry about, though, if we have philosophical misgivings over countenancing entities other than concrete objects.

Russell ([2], [3], *Principia*) had a no-class theory. Notations purporting to refer to classes were so defined, in context, that all such references would disappear on expansion. This result was hailed by some, notably Hans Hahn, as freeing mathematics from platonism, as reconciling mathematics with an exclusively concrete ontology. But this interpretation is wrong. Russell's method eliminates classes, but only by appeal to another realm of equally abstract or universal entities—so-called propositional functions. The phrase 'propositional function' is used ambiguously in *Principia Mathematica*; sometimes it means an open sentence and sometimes it means an attribute. Russell's no-class theory uses propositional functions in this second sense as values of bound variables; so nothing can be claimed for the theory beyond a reduction of certain universals to others, classes to attributes. Such reduction comes to seem pretty idle when we reflect that the underlying theory of attributes itself might better have been interpreted as a theory of classes all along, in conformity with the policy of identifying indiscernibles.

[18] Pp. 90ff, 96ff.

6

By treating predicate letters as variables of quantification we precipitated a torrent of universals against which intuition is powerless. We can no longer see what we are doing, nor where the flood is carrying us. Our precautions against contradictions are *ad hoc* devices, justified only in that, or in so far as, they seem to work.

There is, however, a more restrained way of treating predicate letters as variables of quantification; and it does maintain some semblance of control, some sense of where we are going. The idea underlying this more moderate method is that classes are conceptual in nature and created by man. In the beginning there are only concrete objects, and these may be thought of as the values of the bound variables of the unspoiled theory of quantification. Let us call them *objects of order* 0. The theory of quantification itself, supplemented with any constant extra-logical predicates we like, constitutes a language for talking about concrete objects of order 0; let us call this language L_0. Now the first step of reification of classes is to be limited to classes such that membership in any one of them is equivalent to some condition expressible in L_0; and correspondingly for relations. Let us call these classes and relations *objects of order* 1. So we begin binding predicate letters with the idea that they are to admit objects of order 1 as values; and, as a reminder of this limitation, we attach the exponent '1' to such variables. The language formed by thus extending L_0 will be called L_1; it has two kinds of bound variables, namely the old individual variables and variables with exponent '1'. We may conveniently regard the orders as cumulative, thus reckoning the objects of order 0 as simultaneously of order 1. This means counting the values of '*x*', '*y*', etc. among the values of 'F^1', 'G^1', etc. We can explain 'F^1x' arbitrarily as identifying F^1 with x in case F^1 is an individual.[19]

Now the next step is to reify all further classes of such kind that membership in any one of them is equivalent to some condition expressible in L_1; and similarly for relations. Let us call these classes and relations *objects of order* 2. We extend the term to include also the objects of order 1, in conformity with our cumulative principle. So we begin binding 'F^2', 'G^2', etc., with the idea that they are to take as values objects of order 2.

Continuing thus to L_3, L_4, and so on, we introduce bound variables with ever-increasing exponents, concomitantly admitting increasingly wide ranges of classes and relations as values of our variables. The limit L_ω of this series of cumulative languages—or, what comes to the same thing, the sum of all these languages—is our final logic of classes and relations, under the new procedure.

What we want to do next is specify a theory to much the same effect as L_ω by direct rules, rather than by summation of an infinite series. For purposes of the general theory certain simplifications can be introduced into the above plan. At the stage L_0 there was mention of some initial assortment of extralogical predicates; but the choice of such predicates is relevant only to applications, and can be left out of account in the formal theory in the same spirit in which we pass over the question of the specific nature of the objects of order 0. Furthermore, as noted in another connection at the end of the preceding section, we can omit the polyadic use of bindable variables; and we can rewrite the residual forms 'F^3x', 'G^2F^3', etc. in the preferred notation '$x^0 \in y^3$', '$y^3 \in z^2$', etc. The notation thus becomes identical with that of Essay V, but with exponents added to all variables. There are no restrictions analogous to those of the theory of types: no requirements of consecutiveness, indeed no restrictions on meaningfulness of combinations. Such a combination as '$y^3 \in z^2$' can be retained as meaningful, and even as true for some values of y^3 and z^2, despite the fact that all members of z^2 are of order 1; for, orders being cumulative, y^3 may well be of order 1.

[19] See above, pp. 81f.

Moreover the rules R1-5 of Essay V can be carried over intact, except that restrictions are needed on R2-3. The restriction on R2 is that *the exponent on* β *must not exceed that on* α. The reason is evident: if α takes classes of order m as values and β takes classes of order n as values, then all possible values of β will be included among those of α only if $m \geqq n$. The restriction on R3 is that *'y' and 'x' must bear ascending exponents, and* ϕ *must contain no exponent higher than that on 'x', and none even as high inside of quantifiers.* This restriction reflects the fact that the classes of order $m + 1$ draw their members from order m according to conditions formulable within L_m.

P1 may be retained, but the signs '$\subset$' and '$=$' therein must be redefined now with attention to exponents, as follows: '$x^m \subset y^n$' and '$x^m = y^n$', for each choice of m and n, are abbreviations respectively of:

$$(z^{m-1})(z^{m-1} \in x^m \supset z^{m-1} \in y^n), \qquad (z^{m+1})(x^m \in z^{m+1} \supset y^n \in z^{m+1}).$$

We then also need, for all choices of exponents, the postulate:

$$x = y \supset (x \in z \equiv y \in z) .$$

This theory of classes is closely akin to Weyl's, and comparable in power to Russell's so-called ramified theory of types[20] which was proved consistent by Fitch [2]; but it is far simpler in form than either of those systems. It represents, like those systems, a position of conceptualism as opposed to Platonic realism;[21] it treats classes as constructions rather than discoveries. The kind of reasoning at which it boggles is that to which Poincaré (pp. 43-48) objected under the name of *impredicative definition*, namely, specification of a class by appeal to a realm of objects among which that very class is included. The above restriction on R3 is just a precise formulation of the prohibition of so-called impredicative definition.

If classes are viewed as preëxisting, obviously there is no

[20] Without the axiom of reducibility. See below, p. 127.

[21] See above, pp. 14f. The conceptualist position in the foundations of mathematics is sometimes called *intuitionism*, in a broad sense of the term. Under stricter usage 'intuitionism' refers only to Brouwer and Heyting's special brand of conceptualism, which suspends the law of the excluded middle.

objection to picking one out by a trait which presupposes its existence; for the conceptualist, on the other hand, classes exist only in so far as they admit of ordered generation. This way of keynoting the conceptualist position is indeed vague and metaphorical, and in seeming to infuse logical laws with temporal process it is puzzling and misleading. For a strict formulation of the position, however, free of metaphor, we can point to the above system itself.

Let us see how it is that Russell's paradox is now obstructed. The proof of Russell's paradox consisted in taking the ϕ of R3 as '$\sim(y \in y)$', and afterward taking y as x. Now the first of these steps still goes through, despite the restriction on R3. We get:

(6) $$(\exists x^{n+1})(y^n)[y^n \in x^{n+1} \equiv \sim(y^n \in y^n)]$$

for each n. But the second step, which would lead to the self-contradiction:

(7) $$(\exists x^{n+1})[x^{n+1} \in x^{n+1} \equiv \sim(x^{n+1} \in x^{n+1})],$$

is obstructed. For, the derivation of (7) from (6) by R1, R2, R4, and R5 would, if carried out explicitly, be found to make use of this case of R2:

$$(y^n)[y^n \in x^{n+1} \equiv \sim(y^n \in y^n)] \supset [x^{n+1} \in x^{n+1} \equiv \sim(x^{n+1} \in x^{n+1})].$$

But this case violates the restriction on R2, in that $n + 1$ exceeds n.

Intuitively the situation is as follows. (6), which holds, assures us of the existence, for any n, of the class of non-self-members of order n. But this class is not itself of order n, and hence the question whether it belongs to itself does not issue in paradox.

The conceptualist theory of classes requires no classes to exist beyond those corresponding to expressible conditions of membership. It was remarked in the preceding section that Cantor's theorem would entail the contrary situation; however, his theorem is not forthcoming here. For Cantor's proof appealed to a class h of those members of a class k which are not members of the subclasses of k to which they are correlated.[22] But this way

[22] See p. 92n above.

of specifying h is impredicative, involving as it does a quantification over the subclasses of k, one of which is h itself.

Thus it is that a theorem of classical or semiclassical mathematics goes by the board of conceptualism. The same fate overtakes Cantor's proof of the existence of infinities beyond the denumerable; this theorem is just a corollary, indeed, of the theorem discussed above. So far, good riddance. But obstacles turn out to confront the proofs also of certain more traditional and distinctly more desirable theorems of mathematics; for example, the proof that every bounded class of real numbers has a least bound.

When Russell propounded his ramified theory of types, these limitations led him to add his "axiom of reducibility." But the adding of this axiom, unjustifiable from a conceptualist point of view, has the effect of reinstating the whole platonistic logic of classes. A serious conceptualist will reject the axiom of reducibility as false.[23]

7

The platonist can stomach anything short of contradiction; and when contradiction does appear, he is content to remove it with an *ad hoc* restriction. The conceptualist is more squeamish; he tolerates elementary arithmetic and a good deal more, but he balks at the theory of higher infinities and at parts of the higher theory of real numbers. In a fundamental respect, however, the conceptualist and the platonist are alike: they both assume universals, classes, irreducibly as values of their bound variables. The platonistic class theory of §5 and the conceptualistic class theory of §6 differ only thus: in the platonistic theory the universe of classes is limited grudgingly and minimally by restrictions whose sole purpose is the avoidance of paradox, whereas in the conceptualistic theory the universe of classes is limited cheerfully and drastically in terms of a metaphor of progressive creation. It would be a mistake to suppose that this metaphor really accounts for the classes, or explains them away; for there is no

[23] See my [3].

indication of how the conceptualist's quantification over classes can be paraphrased into any more basic and ontologically more innocent notation. The conceptualist has indeed some justification for feeling that his ground is solider than the platonist's, but his justification is limited to these two points: the universe of classes which he assumes is meagerer than the platonist's, and the principle by which he limits it rests on a metaphor that has some intuitive worth.

The heroic or quixotic position is that of the nominalist, who foreswears quantification over universals, for example, classes, altogether. He remains free to accept the logic of truth functions and quantification and identity, and also any fixed predicates he likes which apply to particulars, or nonuniversals (however these be construed). He can even accept the so-called algebras of classes and relations, in the narrowest sense, and the most rudimentary phases of arithmetic; for these theories can be reconstrued as mere notational variants of the logic of quantification and identity.[24] He can accept laws which contain variables for classes and relations and numbers, as long as the laws are asserted as holding for all values of those variables; for he can treat such laws as schemata, like the laws of truth functions and quantification. But bound variables for classes or relations or numbers, if they occur in existential quantifiers or in universal quantifiers within subordinate clauses, must be renounced by the nominalist in all contexts in which he cannot explain them away by paraphrase. He must renounce them when he needs them.

The nominalist could of course gain full freedom to quantify over numbers if he identified them, by some arbitrary correlation, with the several particulars of his recognized universe—say with the concrete individuals of the physical world. But this expedient has the shortcoming that it cannot guarantee the infinite multiplicity of numbers which classical arithmetic demands. The nominalist has repudiated the infinite universe of universals as a dream world; he is not going to impute infinitude to his universe of particulars unless it happens to be infinite as a

[24] See my [2], pp. 230ff, 239.

matter of objective fact—attested to, say, by the physicist. From a mathematical point of view, indeed, the important opposition of doctrines here is precisely the opposition between unwillingness and willingness to posit, out of hand, an infinite universe. This is a clearer division than that between nominalists and others as ordinarily conceived, for the latter division depends on a none too clear distinction between what qualifies as particular and what counts as universal. In the opposition between conceptualists and platonists, in turn, we have an opposition between those who admit just one degree of infinity and those who admit a Cantorian hierarchy of infinities.

The nominalist, or he who preserves an agnosticism about the infinitude of entities, can still accommodate in a certain indirect way the mathematics of the infinitist—the conceptualist or platonist. Though he cannot believe such mathematics, he *can* formulate the rules of its prosecution.[25] But he would like to show also that whatever service classical mathematics performs for science can in theory be performed equally, if less simply, by really nominalistic methods—unaided by a meaningless mathematics whose mere syntax is nominalistically described. And here he has his work cut out for him. Here he finds the strongest temptation to fall into the more easygoing ways of the conceptualist, who, accepting a conveniently large slice of classical mathematics, needs only to show the dispensability of the theory of higher infinites and portions of real number theory.

Tactically, conceptualism is no doubt the strongest position of the three; for the tired nominalist can lapse into conceptualism and still allay his puritanic conscience with the reflection that he has not quite taken to eating lotus with the platonists.

[25] See above, p. 15.