THE JOURNAL OF PHILOSOPHY
VOLUME LXXXI, NO. 3, MARCH 1984

LOGICISM AND THE ONTOLOGICAL COMMITMENTS OF ARITHMETIC*

FREGE'S published writings on the nature of mathematics (excluding geometry) are organized around two central theses:

(1) Mathematics is really logic.
(2) Mathematics is about distinctively mathematical sorts of objects, for example, the cardinal numbers.

These theses may seem to be uncomfortable passengers in a single boat. Logic is often thought to be unique among the sciences in its lack of a distinctive subject matter, in its "topic-neutrality." How can a part of logic be about a distinctive domain of objects and yet preserve its topic-neutrality? This alone should goad us to clarify the import of (1). But my reasons for pursuing this matter are only secondarily exegetical. Today logicism, the doctrine that (1) struggles to express, is widely regarded as false. I think that it is true. Frege's views are both intrinsically worth understanding and useful as a foil against which to present an alternative story, one which is in part squarely Fregean in spirit and in part contrary to that spirit.

As a gloss on (1) we turn to:

(1.1) All purely mathematical propositions are purely logical propositions.

A proposition is purely mathematical iff all its constituents are mathematical or logical constituents; it is purely logical iff all its constituents are logical. This is still rather vague, but it's clear enough to be puzzling to a student with one term of logic behind him; for a first-order language without identity does not offer us

* Versions of this paper were read at Harvard in 1972 and at Tufts in November of 1981. I have benefited from conversations with George Boolos, Richard Boyd, and Margery Colten.

0022-362X/84/8103/0123$02.70

any sentences that plausibly express purely logical propositions. Admitting identity as a logical notion provides (as W. V. Quine points out[1]) such sentences; but with only these available, (1.1) would be preposterous. The slightest acquaintance with Frege's writings on mathematics eliminates this misunderstanding. Under the rubric "logic" Frege included at least second-order logic with quantification over relations of type 1 and type (0,0).[2] (Indeed identity should be considered a logical notion only because it is the tip of the second-order iceberg—a level 1 relation with a pure second-order definition.) A second-order language has a rich class of pure sentences containing no names or predicate constants, but only variables and the familiar truth-functional and quantificational expressions.

In his published writings on mathematics, his correspondence, and most of his *Nachlass,* Frege repeatedly insisted that numbers were "self-subsistent objects," showing annoyance with his opponents who, he thought, ignored or misunderstood matters of logical form. But what is the real content of thesis (2)? In part it is an assertion about the syntactic form of sentences expressing mathematical propositions; in part it is a claim about the semantic relations that give life to these syntactic forms. It may be spelled out as follows:

(2.1) Numerical terms are designating singular terms; their contribution to determining the truth values of the sentences in which they occur depends on which objects they designate.

(2.2) Numerical predicates stand for level 1 concepts and relations (I'll use 'concept' in its technical Fregean sense); their contribution to determining the truth values of the sentences in which they occur depends on which concepts or relations they designate.

(2.3) Quantification over mathematical objects is to be construed "referentially," not merely "substitutionally."

Put in a nutshell and restricted to cardinal arithmetic:

(2.4) The propositions of pure cardinal arithmetic have the logical forms displayed by their natural regimentation in a first-order formal language.

[1] *The Philosophy of Logic* (New York: Prentice-Hall, 1970), p. 61.

[2] I follow the usual notation for logical types: '0' is a type symbol representing the type of objects; if $\sigma_1, \ldots, \sigma_n$ are type symbols, then $\ulcorner(\sigma_1, \ldots, \sigma_n)\urcorner$ is a type symbol representing the type of n-place relations whose ith component is of type σ_i for $1 \leq i \leq n$; $\ulcorner n + 1\urcorner$ is $\ulcorner(n)\urcorner$; level(0) = 0; level$((\sigma_1, \ldots, \sigma_n))$ = max(level$(\sigma_1), \ldots,$ level$(\sigma_n))$ + 1. Types classify linguistic expressions as well as what they stand for. '$\forall$' is typically ambiguous, but in a well-formed context is disambiguated: binding a variable of type σ, it represents the universal quantifier of type $((\sigma))$.

Once this much is accepted, further familiar metaphysical theses become close to irresistible:

(3) Numbers (what numerical terms designate, what numerical predicates apply to, what number-theoretic quantifiers range over) exist outside of space and time; they are in no sense created by us, by thought, or by "mathematical activity" (whatever that may be).

This is all prima facie plausible. After all, sciences have subject matters: ichthyology is about fish, ecology about eco-systems, etc. These subject matters are kinds or types of objects of a reasonably non-ad-hoc sort. The mathematical sciences seem to be no exception. It would sound silly to deny that a book on number theory is about numbers. This is how mathematicians and nonmathematicians alike speak. So once we have learned some logic and start seeing the logical form of statements about people, planets, positrons, and nations through the goggles of the first-order predicate calculus, and all looks clear and basically right, and we come across syntactically similar discourse in Mathematics departments and mathematics texts, it's only natural, for generalizing creatures like ourselves, to try on these goggles in this land, too. Within certain limits, to be discussed momentarily, things look fine: mathematical discourse *can* be regimented into first-order languages, and mathematical arguments can be formalized in any complete formalization of first-order logic.[3] All this is built into doctrine (2), and it all seems right. When other considerations brought Frege's research project to grief, he seems to have been quite sensible to hold to his slogan, "Numbers are objects," as a fixed point, a beacon of light in the confusing darkness:

> The prime problem of arithmetic may be taken to be the problem: How do we apprehend logical objects, in particular numbers? What justifies us in recognizing numbers as objects? Even if this problem is not yet solved to the extent that I believed it was . . . , nonetheless I do not doubt that the way to a solution has been found.[4]

Several passages in the *Nachlass* suggest that late in his life Frege was inclined to abandon this fixed point. I contend that this is the move that should be made. It is natural, harmless—at least when one is not doing philosophy—and even helpful, to think and speak

[3] Perhaps Category theory is an exception.

[4] See the Appendix of Volume II of *The Basic Laws of Arithmetic,* in Peter Geach and Max Black, *Translations from the Philosophical Writings of Gottlob Frege* (Oxford: Blackwell 2d ed., 1970), p. 244. Henceforth this book will be cited as "Geach & Black."

as if there were numbers, to pretend to posit numbers, to pretend that (2.1)-(2.3) are true. Speaking Wittgensteinese, to engage in this pretense shall be called "accepting the mathematical-object picture." Accepting a picture is rather like employing a model in the ordinary nontechnical sense. In this sense, a model is not a theory, even when it's employed for theoretical purposes, e.g., to help one understand or apply a theory. A model is useful or informative because of what Mary Hesse called a "positive analogy" between it, a well-understood portion or aspect of reality, and that ill-understood portion of reality which it models. Analogies are analogies in certain, and not in all, respects. To fully understand a model one must see "where" the sustaining positive analogy runs out. The same goes for philosophical pictures. Frege and those who follow him in accepting (2.1)-(2.4) convert a benign picture into a false theory. This is similar to what is done by the misguided physics student who pushes a physical model beyond its sustaining positive analogy, e.g., one who, on learning that an atom is like a little solar system, thinks that the electrons are tiny chunks of rock revolving around the nucleus in fixed, well-defined coplanar orbits.

Of course, science does not rest content with mere models. We at least *seek* a degree of understanding which permits a model-independent formulation of the sustaining positive analogy and, thereby, an explanation of the value the model has had for us. Analogously in this case: if (2.1)-(2.4) are false, we owe an account of why it's useful, natural, and harmless (outside of philosophy) to pretend that they are true.

I. WHAT IS WRONG WITH THE MATHEMATICAL-OBJECT THEORY?

If we accept the mathematical-object theory, we then owe ourselves an account of the nature of reference to mathematical objects. This is not to say that we owe ourselves necessary and sufficient conditions, explicitly formulated in "more basic terms," defining what it is for an expression, or the user of an expression, to refer to a mathematical object; such a demand is surely scientifically utopian. But the phenomena of word reference and of speaker reference, and our referential abilities themselves, have a microstructure about which we can give an informative account. I'm urging that the availability of such an account constrains our choice of an ontology. (Of course it doesn't uniquely determine one.) The challenge to the mathematical-object theorist is: Tell us about the microstructure of reference to, e.g., cardinal numbers. In what does our ability to refer to such objects consist? What are the facts about our linguistic practice by virtue of which expressions in our language designate such objects and the concepts under which they fall or fail to fall?

If no plausible answers are available, we should reject the theory that invites these questions. We should conclude that in asking these questions we were overstepping the positive analogy that sustains a model, applying to a picture a weight that only a theory could bear. But rejecting the theory is not the end of the matter. We now owe another account of the logical form of mathematical statements; I'll sketch a way to make good this debt.

Our paradigm cases of reference are the everyday cases of referring to persons, places, "medium-sized pieces of dry goods," macroscopic events, and natural and artifact kinds. In the microstructure of these cases, as well as in the more recherché cases of reference to theoretical entities in scientific discourse, causal relations and causal facts loom large. Recognizing this accomplishes little, since causation is a very loose, amorphous, and pragmatic concept. But whatever the links between ourselves, our practices and abilities, and objects like Jupiter, Exxon, tyrannosaurus rex, and positrons, by virtue of which we and our words refer to them, these links are going to have to be rather different from any such links between ourselves, our practices, our abilities, and the number 1. Numbers are so pure, so unstained by the cement of the universe, that reference to them and their ilk seems quite *sui generis.*

Frege appreciated this point. He called numbers "logical objects" not merely because he thought them to be objects in his technical "logical" sense, but because he thought that the microstructure of reference to numbers involved logic in a way that made it distinctively different from the microstructure of reference in what I've called the "paradigm" cases. *The Foundations of Arithmetic* may be read as an attempt to exhibit the peculiarly logical aspect of reference to numbers. Frege's question:

> How then are numbers given to us, if we cannot have any idea or intuition of them?

is his rather nineteenth-century formulation of our challenge to the mathematical-object theory. Discussions of reference before this century often focused on the relation of thought, rather than language, to reality, invoking perceptual intuition as a paradigm case. In saying:

> We can form no idea of the number [4] whether as a self-subsistent object or as a property in an external thing, because it is not in fact either anything sensible or a property of an external thing.[5]

[5] *The Foundations of Arithmetic,* J. L. Austin, trans. (New York: Harper 1960), p. 70.

Frege acknowledges the extent to which reference to numbers differs from the paradigm cases of reference. He contends that if this difference makes reference to numbers seem problematic then we're conceiving of reference too atomistically.

I think it's not overly anachronistic to restate what Frege said in *The Foundation of Arithmetic* using the terminology of "On Sense and Reference," published three years later. In that terminology, for an object to be in my referential ken is for me to grasp a sense that "presents" that object. To grasp a sense is, at least in part, a matter of being readily able to grasp thoughts of which that sense is a constituent.

> That we can form no idea of its content [referent] is therefore no reason for denying all meaning [sense] to a word, or for excluding it from our vocabulary. We are indeed only imposed on by the opposite view because we will, when asking for the meaning of a word, consider it in isolation, which leads us to accept an idea as the meaning. Accordingly, any word for which we can find no corresponding mental picture appears to have no content. But we ought always to keep before our eyes a complete proposition [sentence]. Only in a proposition have the words really a meaning. . . . It is enough if the proposition taken as a whole has a sense; it is this that confers on its parts also their content [sense] (*ibid.,* 71).

The scope of these remarks is unclear. If they are to apply to all cases of reference, then Frege would be committed to the claim that all thought-grasping is prior to all grasping of the senses of subsentential constituents and so prior to all reference; this is to say that in all cases the microstructure of thought-grasping itself involves *no* facts about reference and thus can be explicated without adverting to any cases of sense-grasping for subsentential constituents. This is implausible. Fortunately for Frege's project, such generality is not required. The thoughts at issue are mathematical propositions. So it would be necessary only that one be able to characterize the grasping of *these* thoughts without adverting to any cases of reference *to mathematical objects.* If this less ambitious project succeeds, then Frege could offer his "odd sort of definition" of number words as an account of what it is to grasp the senses of such words and, thus, as an answer to our challenge to the mathematical-object theory. Frege is saying: "If I explain what it is to grasp certain thoughts and how number words figure in expressing these thoughts, I will have shown how we come to grasp the senses of number words and thus how numbers enter our referential ken."

This endeavor presupposes a curious and important Fregean doctrine: that thoughts have polymorphous composition. I have discussed this at length elsewhere[6] and will briefly review the matter here. Consider these examples.

(i) There is a moon of Jupiter.
(ii) There are exactly four moons of Jupiter.
(iii) the number of moons of Jupiter = 4

Frege drew attention to the analogy between (i) and (ii): both may be taken as "statements about concepts," that is, as quantificational sentences in which the predicate 'is a moon of Jupiter' completes a quantificational expression. These expressions stand for object-quantifiers—level 2 functions of a certain sort. In (i) the quantifier in question is the familiar existential object-quantifier. In (ii) it is a less familiar cardinality object-quantifier. (ii) may be rephrased as "There are exactly four x such that x is a moon of Jupiter" and parsed as:

(ii′) $(\exists 4x)$ (x is a moon of Jupiter)

Notice that in (ii′) the occurrence of '4' is syncategorematic, playing no referential role; it is not even syntactically a singular term: '$\exists 4$' is a single syntactic unit. As is well known, the expression '$(\exists 4x)$' may itself be defined within first-order logic with identity; in this case (ii′) becomes:

$$\text{(ii}''\text{)}\ (\exists x_1)(\exists x_2)(\exists x_3)(\exists x_4)\Big(\bigwedge_{1 \leqslant i \leqslant 4} x_i \text{ is a moon of Jupiter } \&\ \bigwedge_{1 \leq i < j \leq 4} x_i \neq x_j\ \&\ (\forall y)\,(y \text{ is a moon of Jupiter} \supset \bigvee_{1 \leqslant i \leqslant 4} y = x_i)\Big)$$

(ii), (ii′), and (ii″) all express the same thought; (ii″) shows us how one could grasp that thought with a prior grasp only of the senses of '=' and of quantificational and truth-functional expressions, and so without having the number 4 itself in one's referential ken. But, and here's the crucial point, (ii) and (iii) also express the same thought. (ii) and (iii) differ in the way they display the composition of that thought, but according to Frege, one thought is not composed out of a unique set of atomic senses in a unique way. The problem that remains is this: given that we can explicate what it is to grasp thoughts like those expressed by sentence (ii), how do we explicate what it is to grasp the sense of '4' or 'the number of moons of Jupiter'?

At this point, let me put some important cards on the table. I am accepting the Fregean ontology of objects and functions of various levels. I've heard Frege's views on objects and functions disparaged

[6] "The Composition of Fregean Thoughts," *Philosophical Studies*, XLI, 2 (March 1982): 161-178.

as "the least plausible of his views," solely, I think, because of the "paradox" of the concept of a horse. But I think that these doctrines are among the most plausible and important things Frege has offered us, that this so-called "paradox" is not a paradox, that Frege's request that his reader "not begrudge a pinch of salt" is entirely reasonable, and that failure to appreciate the fundamental difference between saturated and unsaturated or predicative entities is as much a potential source of confusion today as it was one hundred years ago.[7] Why is this important? Because the Fregean notion of a predicative entity seems necessary for the intelligibility of higher-order logic as it is usually interpreted.

But is second-order logic intelligible? Quine, for example, is in places unwilling to suppose so, for he holds to Mill's account of the referential role of a predicate: a predicate multiply denotes any and all objects to which it applies; and there is no need to posit a further "predicative" entity to which a predicate stands in some further referential relation, one somehow more basic than (multiple) denotation. Unless we posit such further entities, second-order variables are without values, and quantificational expressions binding such variables can't be interpreted referentially. In a second-order language in which 'x' and 'X' are variables of types 0 and 1, respectively, 'Xx' is a formula, and 'X' must range over predicative entities. To regard 'Xx' as abbreviating '$x \in X$' is equivalent to insisting that the values of 'X' be objects; in this case we'd no longer have a second-order language, but rather a first-order language involving two styles of variables of logical type 0 and a binary predicate '$\in$' which, in Fregean terms, stands for a relation of level 1.

Quine has urged us to take first-order logic as the measure of all things, or at least of all ontologies. I think this would be a mistake, even if Mill's notion of multiple denotation were adequate for an account of the semantics of predicates in first-order languages. I

[7] David Wiggins writes "It would be good some day to show that this kind of talk is fully compatible with nominalism in that reasonable acceptation of the term in which both Aristotle and Leibniz are to be reckoned nominalists" [*Sameness and Substance* (Cambridge, Mass.: Harvard, 1980)]. This "thin" understanding of the reference of incomplete expressions is developed by Michael Dummett in *Frege: Philosophy of Language* (London: Duckworth, 1973; parenthetical page references to Dummett will be to this book) and by Montgomery Furth in "Two Types of Denotation", in Nicholas Rescher, ed., *Studies in Logic* (Oxford: APQ Monograph Series, 1968). It is contested in Edwin Martin, "Frege's Problems with 'the Concept of *Horse*'" *Critica*, v (September 1971): 45-64. In the original twelfth-century sense of the term, the doctrine that predicates stand for concepts, where concepts are in no sense mental entities, is realism, not nominalism.

shall not directly argue this point.[8] Suffice to say: brand higher-order logic as unintelligible if you will, but don't conflate it with set theory.[9]

The impulse for this conflation has several sources. First, there is the bad example of Russell and Whitehead. Furthermore, I think there is a deep connection between higher-order logic and set theory which remains ill understood. But the major source for this conflation seems to lie in another conflation: that of truth with truth in a model.

Following somewhat standard usage (e.g., that of C. C. Chang and H. Jerome Kiesler in *Model Theory*[10]), I'll regard a model as a set of a certain sort, associated with a certain uninterpreted language: an ordered pair (U, V), where U is a non-empty set and V is a function on the nonlogical lexicon of the language assigning these lexical items to appropriate members of U or sets constructed from U. Sentences of that language may be related to such a model by the "true-in" relation. V is often said to "interpret" the language—a harmless use of the word, provided we keep in mind that V does not associate anything like Fregean senses with sentences or lexical items. Truth, however, is a property of sentences in an interpreted language, where 'interpreted' as used here expresses a notion more robust than its thin-blooded model-theoretic namesake. Truth in a model is interesting because it provides a transparent and mathematically tractable model—in the "ordinary" sense discussed previously—of the less tractable notion of truth. This modeling relation makes possible a first-order (set-theoretic) definition of the prima facie second-order notions of implication and validity.

Conflation of truth with truth in a model is encouraged by the following phenomena: sometimes an interpreted language is concerned with "a restricted portion of reality"; if this restricted domain constitutes a set, then we may collapse truth into truth in an "intended" model. Thus, for the sake of argument, suppose there are natural numbers and, furthermore, there is a set of all of them;

[8] In *Nominalism and Realism* (New York: Cambridge, 1978) D. M. Armstrong calls Quine's view on this matter "Ostrich or Cloak and Dagger Nominalism"; in several places he rejects it because we must posit universals to explain why, for example, a particular piece of paper is white (30). The nature of this "why" question is unclear: if "There is a property of whiteness and this paper just has it" counts as a permissible answer, then I'd suppose that the question was illegitimate; if the answer could be in terms of the physical microstructure of the paper, it's unclear that reification of universals is involved.

[9] For further discussion, see George S. Boolos, "On Second-order Logic," this JOURNAL, LXXII, 16 (Sept. 18, 1975): 509-527.

[10] New York: American Elsevier, 1974.

we then have the so-called "intended" model for arithmetic; and truth in an interpreted language (in which all quantifiers are restricted by the predicate 'is a natural number' and all designators designate natural numbers) collapses to truth in this model. But set-theoretic discourse, as usually understood, provides an example for which this cannot happen. One thing the beginning student of set theory must get clear about is the enormous difference between the truth of sentences in the interpreted language of set theory and truth in some model for the disinterpreted skeleton of that language.[11]

How does conflation of truth with truth in a model encourage conflation of higher-order logic with set theory? When we define truth in a model for an uninterpreted second-order language, the second-order variables may be taken to range over sets instead of over predicative entities. This is possible because the domain of a model is a set, and so the predicative entities do have extensions when restricted to such a domain. But to go from this to thinking that in an interpreted language second-order variables range over sets would be overstepping the positive analogy in virtue of which truth in a model models truth.

In our discussion of (ii) we came across a cardinality object-quantifier which in (ii′) is represented by '$(\exists 4)$'. In general, a cardinality object-quantifier Q is a type 2 concept such that for any type 1 concepts X and Y:

$$(Qx)Xx \supset ((Qx)Yx \equiv (Q_{\mathrm{E}}x)(Xx, Yx))$$

where '$(Q_{\mathrm{E}}\nu)(\phi, \psi)$' means that the ν's such that ϕ are equinumerous with the ν's such that ψ.

In (iii) we find a form of term construction which merits attention: the syntactic role of the phrase 'the number of' in cases like (iii) is to convert a formula into a singular term, in the process binding at most one variable occurring in that formula. We represent this using '#' and the rule:

If ϕ is a formula and ν is a variable, then $(\#\nu)\phi$ is a singular term.

[11] Even here mathematical usage can be misleading, for some set-theorists do not view models as sets, but refer to the universe of all sets, the constructible universe, the core universe, etc., as "inner models." A distinction between "concrete" and "abstract" models has had some currency as a device for avoiding confusion. More commonly, it will be said that the latter sorts of models are "proper classes." This terminology is acceptable, provided that we do not construe proper classes as objects in the Fregean sense; proper classes are essentially predicative entities. If we're told that 'X' is a class variable and '$x \in X$' is a formula, we should automatically read it as an unfortunate replacement for 'Xx'. I trust that all this is boring pedantry to the set-theoretical sophisticate; but for others these matters are a mine field of possible

As Frege emphasized, the semantics behind this construction relates it to equinumerosity, as shown by the truth of all instances of this schema:

$$\text{(A)} \qquad (\#\nu)\phi = (\#\nu)\psi \equiv (Q_{\mathrm{E}}\nu)(\phi, \psi)$$

To facilitate description of the semantics of this construction, let's introduce some technical notions.

By a *numberer* I mean any type 2 function F carrying type 1 concepts to objects, such that, for all such concepts X and Y:

$$F(X) = (F(Y) \equiv (Q_{\mathrm{E}}x)(Xx, Yx)$$ [12]

A *representor* is a type 3 function G carrying cardinality object-quantifiers to objects, such that, for cardinality quantifiers Q and Q',

$$G(Q) = G(Q') \equiv (\forall X)((Qx)Xx \equiv (Q'x)Xx)$$

We ignore arguments that are not cardinality quantifiers. Keep in mind that the values of numberers and representors may be any objects. Clearly numberers and representors come in pairs: a numberer F is paired with a representor G iff F assigns to all concepts falling under a given cardinality quantifier what G assigns to that quantifier.

We can now formulate the key commitment of the mathematical-object theory applied to cardinal arithmetic: among all the numberers one is special; it is what we'll call the *standard numberer,* assigning to a type 1 concept the number of objects falling under that concept. This is what the phrase 'the number of' represents: if 'ϕ' has at most 'ν' free and stands for the type 1 concept X, then '$(\#\nu)\phi$' stands for the value of X under the standard numberer. Paired with the standard numberer is the *standard representor,* which assigns to each cardinality object-quantifier a special sort of object intrinsically, internally, and just plain *specially* related to that quantifier. The problem with which we left Frege several pages back was this: Suppose that the cardinality quantifiers are in our referential ken; and suppose that the microstructure of this fact can be explicated without adverting to a primitive ability to refer to numbers; how can we pass from this to an account of what it is to have the numbers themselves in our referential ken? As Frege puts it, "How

---

confusions through which many standard texts lead the student with little guidance or protection.

[12] Using the familiar bound-variable notation to express the way in which higher-level functions complete lower-level ones, '$F(X)$' abbreviates '$(Fx)Xx$', and '$G(Q)$' abbreviates '$(GX)((Qx)Xx)$'.

can we get from these concepts to the numbers of arithmetic in a way that cannot be faulted?"[13]

To better appreciate the challenge facing the mathematical-object theory, we will consider the following fantasy. Suppose Adam speaks a language just like English, except in the following respect: in his language '4' designates 5, '5' designates 4, 'successor' stands for a function differing from the successor function only in assigning 3 to 5, 5 to 4, and 4 to 6; 'less than' stands for a relation like less than, except that 5 bears it to 4 and not the reverse; references to other arithmetic relations are similarly skewed; and finally, 'the number of' stands for a nonstandard number differing from the standard numberer only in assigning 5 to concepts under which four objects fall, and 4 to concepts under which five objects fall. There is no systematic divergence between the sentences we accept and those which Adam accepts. Indeed, Adam could dwell among us and no one, including Adam, could discover that his language differed from ours. Now is this case really possible?

I contend that it is not possible. My reasons are not verificationist, but rather are based on the belief that reference supervenes on more basic physical, psychological, and social facts, relations, and phenomena. What could it be about the world and the roles played by Adam and ourselves within it in virtue of which he refers to 5 by '4' and we do not? What conceivable microstructure could this fact have? How can I educate a child to make sure he gets reference to numbers right? In what way did my arithmetic education differ from Adam's so as to leave us speaking different languages? Invoking mathematical intuition here would be like saying "Human minds have access to a fifth dimension in which the cardinal numbers are strung out like perfect pearls, and our mental fingers can just point to them in order to fix the references of our number words"—a charming metaphor perhaps, but not even a start at an answer.

It is important to see the relatively narrow scope of this argument. In certain cases the possibility of systematic and virtually undetectable differences between the reference relations underlying remarkably similar languages may be a genuine possibility. But if such cases strike us as genuine possibilities this is because of our inarticulate understanding of what the microstructure of reference in these cases is, an understanding which if pressed would suggest

[13] Letter to Ludwig Darmstaedter, July 26, 1919, in *Posthumous Writings,* Hans Hermes, Friedrich Kambartel, and Friedrich Kaulback, eds. (Chicago: University Press, 1979), p. 257. This book will be cited hereafter as "Hermes *et al.*"

candidate physical, psychological, or social differences to ground these referential differences.[14]

Is the mathematical-object theory committed to the possibility of the case of Adam? If so, I've argued that it cannot pay its debt to the theory of reference, and so must be rejected. But perhaps it is not so committed.

In that case, its defenders still have some serious explaining to do. For now we have another apparently occult fact: that we all end up speaking languages in which, out of all possible numberers, the phrase 'the number of' stands for the standard numberer. Why is the standard numberer a "reference magnet" which "draws" reference by that phrase, in a way in which its nonstandard competitors cannot?

I suspect that Frege at least dimly appreciated the strangeness of this so-called fact, and made a well-meaning though inadequate gesture toward explaining it. He took the key to this matter to be this: instances of (A) are not merely true, not merely analytically true; the sentences flanking the 'iff' '≡' actually express the same

[14] Quine gives a generalized version of this argument to defend his notorious "inscrutability of reference" thesis. Hilary Putnam has recently offered, in *Reason, Truth and History* (New York: Cambridge, 1981), a maximally general version of this argument, intended to dislodge a picture (or theory?) which he labels "Metaphysical Realism." Though this is not the place for extended discussion, I can't resist some further comments.

There is a wide variety of relations between words of English and other entities, which, when treated in a truth definition as if they were the reference relation for English, assign to all sentences their genuine truth conditions in all possible worlds. The distribution of truth values to sentences in possible worlds does not implicitly define reference. Putnam's conclusion is: to maintain that one of these relations is genuine reference and the others are not is to accord to one of these relations an "occult" status; there are no facts about our linguistic practices and our general relations to the world which could form a basis for saying that one of these relations is special. I contend that there are such facts; for example, when a child says "I want a cherry," she's likely to grab for a nearby cherry and not a nearby cat. Now the child may be grabbing* for a cat, where grabbing* is the relation assigned by one of our pseudo-reference relations to the word 'grabbing'. But this is irrelevant. Of course, in claiming that this is irrelevant, I accord grabbing a special status which I deny to grabbing*. Putnam would, no doubt, reply that this is to make grabbing occult.

According to the realist, languages evolve so as to "carve the world at its joints"—at least at those joints which we can detect and which are of interest to us. Unless one denies that the joints are really there, refers to and refers* to (like grabs and grabs*) are relations that are not on a par. Of course an anti-realist would maintain that the joints aren't really there: to make a distinction between potential carvings that are "at the joints" and those which are not is merely to accord the former an occult status. At least on this point we have a stand-off.

Here, though, is where the case of Adam is different. In that case, there really are no facts analogous to the fact about children, cherries, and 'cherry' just considered. If the case of Adam were possible, reference to mathematical objects would be occult in a way in which reference to cherries, grabbing cherries, etc., is not.

thoughts. In section 66 of *The Foundations of Arithmetic* he suggests that these identities determine a unique assignment of senses to all terms of the form '$(\#\nu)\phi$'; since sense determines reference, these identities determine the standard numberer to be that numberer thereby associated with '#'. In several places he compares this with the way in which a system of simultaneous equations in several unknowns may determine a unique solution.

But do all possible instances of this schema:

(B) the thought that $(x_{\phi} = x_{\psi})$ = the thought that $(Q_{\mathrm{E}}\nu)(\phi, \psi)$

where ϕ and ψ have only ν free, constitute a system solved by a unique assignment of senses to the variables of the form 'x_{ϕ}'? If so, the sense of '$(\#\nu)\phi$' would be the value of 'x_{ϕ}'. But there is no reason to expect such a solution to be unique. Just as truth does not implicitly define reference, so the totality of referents of the right-hand side of instances of (B) do not implicitly define the senses of terms of the form '$(\#\nu)\phi$'. Granted that all thoughts of that form are polymorphously composed, there is no reason to think that there is only one way of composing them as identities.

This objection to Frege might be thought irrelevant, for he seems to go on in *The Foundations* to reject the suggestion of section 66 anyway. He points out that he has not shown how senses can be determined for all equations involving terms of the form '$(\#\nu)\phi$'. The previous discussion has not reconstructed our grasp of the thought expressed by sentences like:

(iv) The number of moons of Jupiter = England.

Here Frege goes on to make his fatal move: he introduces extensions of concepts, or, more generally, *courses-of-values* of functions.

But what follows section 66 does not show that the objection to section 66 is irrelevant. Introducing courses-of-values pushes the problem back from numbers to courses-of-values: What fixes the senses of course-of-value terms? Now there is nothing wrong with pushing a problem back. This move has the virtue of significant generalization, assimilating the abstraction of number to the more general phenomenon of the abstraction of courses-of-values. But it does not solve the problem under consideration. In *The Basic Laws of Arithmetic* Frege avoids the set-theoretic analog of the problems posed by (iv) by restricting himself to a language in which all singular terms are course-of-value abstracts. He then retells a familiar story: all instances of this schematic equation are true:

the thought that $(\acute{\varepsilon}.f(\varepsilon) = \acute{\varepsilon}.g(\varepsilon))$ = the thought that $(\forall\varepsilon)(f(\varepsilon) = g(\varepsilon))$

these simultaneous identities uniquely determine the senses of all course-of-value abstraction terms. So the basic act is still that of section 66 of *The Foundations,* replayed on a wider stage. Frege's cryptic remark in *The Foundations* about attaching "no decisive importance" to the introduction of extensions into his story suggests that he was aware of this.

Here Frege's positive account ends. In "On Function and Concept" he says:

> The possibility of regarding the equality holding generally between values of two functions as a (particular) equality, viz. an equality between ranges of values, is, I think, indemonstrable; it must be taken as a fundamental law of logic (Geach & Black, 26).

Calling a very general axiom of abstraction "a fundamental law of logic" does nothing to explicate our grasp of the senses of expressions that stand for the abstracta, and thus does nothing to explain our puzzling success at referring to such objects. In *The Basic Laws,* he remarks:

> . . . we said: If a (first-level) function (of one argument) and another function are such as always to have the same value for the same argument, then we may say instead that the range of values of the first is the same as that of the second. We are then recognizing something in common to the two functions, and we call this the value-range of the first function and also the value-range of the second function. We must regard it as a fundamental law of logic that we are justified in thus recognizing something common to both (Geach & Black, 179).

Then, in a revealing footnote, he appeals to an ethereal sort of ostention:

> In general, we must not regard the stipulations in Vol. i, with regard to primitive signs, as definitions. Only what is logically complex can be defined; what is simple can only be pointed to (180).

This sounds like an appeal to Kantian pure intuition—a desperate move, given Frege's emphasis elsewhere on the difference between laws of logic and what intuition offers us. In any case, if such ostention is available for courses-of-values, it should also be available directly for the cardinal numbers themselves. I suspect that this explains why Russell's paradox seemed so devastating, not just to Frege's set-theoretic approach in *The Basic Laws,* but to the very thesis that cardinal numbers are objects. Russell's paradox shakes one's faith in all such abstraction axioms and in the veracity of whatever sort of a priori intuition might make these axioms attractive.[15]

We've landed on an interesting tangent which deserves exploration. The abstraction of courses-of-values—or, for our purposes, of the extensions of type 1 concepts—is actually rather different from the abstraction of cardinal numbers of type 1 concepts.

Let an extensor H be a type 2 function assigning for each type 1 concept an object such that, for any type 1 concepts X and Y:

$$H(X) = H(Y) \text{ iff } (\forall x)(Xx \equiv Yx)$$

Until he received his famous letter from Russell, Frege believed that there was a standard extensor, represented by the phrase 'the set of x such that', which assigned to a type 1 concept its extension. What Russell showed was not only that there is no standard extensor, but that there are no extensors at all. That is, not merely is:

$$(\forall X)(\exists x)(\forall y)(y \in x \equiv Xy)$$

false, where 'ϵ' is supposed to express a determinate sense and stand for a determinate relation, but the following formula involving branching quantifiers[16] is unsatisfiable:

$$\begin{matrix} (\forall X)(\exists x) \\ (\forall Y)(\exists y) \end{matrix} \; (x = y \equiv (\forall z)(Xz \equiv Yz))$$

On the other hand, no analogous paradoxes surround the existence of numberers and representors!

$$\begin{matrix} (\forall X)(\exists x) \\ (\forall Y)(\exists y) \end{matrix} \; (x = y \equiv (Q_E z)(Xz, Yz))$$

is satisfiable; in fact, if we accept standard set theory, it is true. (Ultimately, this disanalogy is not so great as it might seem. This tangent will be pursued elsewhere.)

Returning to the main thread, we've seen that the inconsistency

[15] In a piece dated 1906, Frege says of his axiom of abstraction: "Of course it isn't as self-evident as one would wish for a law of logic. And if it were possible for there to be doubts previously, these doubts have been reinforced by the shock the law has sustained from Russell's paradox" (Hermes *et al.*, 256/7).

[16] The appropriateness of branching quantifiers here deserves notice. Similarly, the Burali-Forti paradox shows the unsatisfiability of:

$$\begin{matrix} (\forall X^{(0,0)})\,(\exists x) \\ (\forall Y^{(0,0)})\,(\exists y) \end{matrix} \; (x = y \equiv (\mathcal{I}uv)(X^{(0,0)}uv,\ Y^{(0,0)}uv))$$

where $(\mathcal{I}uv)(\phi, \psi)$ express the order isomorphism of the relations defined by ϕ and ψ in the variables u and v.

of *The Basic Laws* was really only a minor flaw in the Fregean project. Its fundamental flaw was its inability to account for the way in which the senses of number terms are determined. It leaves the reference-magnetic nature of the standard numberer a mystery.

So the mathematical-object theorist must choose between Scylla, which is the possibility of Adam's case, and the Charybdis of reference magnetism. But even before we unearthed this dilemma, the doctrine should have seemed suspect. For what is a number? It is an object that canonically represents a cardinality quantifier. This notion of canonicity must be very strong—for the correspondence between numberers and cardinality object-quantifiers provided by the standard representor is not a matter of convention; rather it is the nature of a cardinal number to be intimately related to a particular cardinality object-quantifier. The number 4 is an object whose entire essence consists in the fact that it represents a certain quantifier. But this fact is merely the fact that the standard representor assigns that quantifier to the number 4. And what's so special about the standard numberer? Only that it matches up cardinality object-quantifiers with the right objects—those which intrinsically represent them. The problem is not with essentialism *per se*. (Perhaps there was a sperm and an egg such that my essence is determined by the fact that I developed from them. But it's not essential to the relation of "developing from" that it relates that sperm and egg to me; here we have no circle; the relation stands on its own in a way in which the standard representor does not.) The essence of a cardinal number is parasitic on the standard representor, and the standard representor's canonicity is ungrounded or, at best, parasitic on the cardinal numbers themselves. This small circle, once noticed, should make talk about a standard representor, and thereby talk about its values, the cardinal numbers, sound like myth making; the requisite sort of canonicity seems unintelligible.

II. CODING FICTIONALISM

Frege's *Nachlass* contains several rather despairing remarks about his efforts to understand reference to numbers. In a letter to one Ludwig Darmstaedter written in 1919, he introduces cardinality object-quantifiers, considers the problem of how they could be represented by objects in an appropriately canonical way, and then seems to consider seriously the possibility that such canonical representation may be unnecessary:

> Since a statement about numbers based on counting contains an assertion about a concept, in a logically perfect language a sentence used to make such a statement must contain two parts, first a sign for the concept about which the statement is made, and secondly a sign for a sec-

> ond level concept [a cardinality object-quantifier]. These second level concepts form a series and there is a rule in accordance with which, if one of these concepts is given, we can specify the next. But still we do not have the numbers of arithmetic; we do not have objects, but concepts. How do we get from these concepts to the numbers of arithmetic in a way that cannot be faulted? or are there simply no numbers in arithmetic? Could the numerals help form signs for these second level concepts and yet not be signs in their own right? (Hermes *et al.*, 263/4)

At the end of this quote Frege has, I think, stumbled onto the right track. In *The Foundations* he has already recognized that all purely arithmetic propositions expressible in the quantifier-free fragment of the first-order language of arithmetic are also expressed by sentences in which all numerals are pressed into syncategorematic roles as parts of expressions denoting numerical object-quantifiers; for example, (v) and (vi) express the same proposition

(v) $5 + 7 = 12$

(vi) $(\forall X)(\forall Y)(((\exists 5x)Xx \mathbin{\&} (\exists 7x)Yx \mathbin{\&} \sim(\exists x)(Xx \mathbin{\&} Yx)) \supset (\exists 12x)(Xx \vee Yx))$

Though it may look forbidding, (vi) expresses what "the person on the street" would think (v) expresses: if one takes five objects and then another seven distinct objects, one has twelve objects in all.[17] (vi) is a logical truth in pure second-order logic; according to the mathematical-object theory (v) involves reference to distinctively mathematical objects; but, for Frege in *The Foundations,* (v) and (vi) exhibit the polymorphous composition of a single thought.

On the other hand, the conclusion of the previous section (p. 139) suggests that (v) does not represent a distinct mode of composition of the thought (v) and (vi) express, but rather that (v) merely abbreviates (vi), which represents the sole analysis of that thought. I'll next consider two ways in which this view could be developed; this discussion will lead us to confront an ambiguity in the notion of logical form.

One might suggest that numerical terms in (v) stand for numerical quantifiers, that numbers are such quantifiers. This suggestion faces an immediate Fregean objection: Singular terms *cannot* stand for quantifiers; modulo the well-known grains of salt, reference to object-quantifiers occurs only in predicative constructions in which the expression standing for an object-quantifier either completes a level 3 expression or is completed by a level 1 expression.

[17] George Boolos has pointed out to me that (vi) does not express the ordinal content of (v).

Does this response make too much of type-theoretic distinctions? Suppose 'P' and 'Q' are type 3 and type 2 constant expressions; in the usual variable-binding notation, '$(PX)((Qx)Xx)$', where 'x' and 'X' are type 0 and type 1 variables, is the result of predicating 'P' of 'Q'. This could reasonably be abbreviated as '$P(Q)$'. In a language without level 0 or level 1 constants, we *could* take '$P(Q)$' as an atomic sentence, pretend that 'P' and 'Q' are of types 1 and 0, and no harm would be done. The language of pure elementary number theory is such a language.

This may seem to make the assignment of logical levels to sentential constituents language-relative, suggesting that one man's concept may be another man's object. Michael Dummett is troubled by the possibility that:

> . . . there is no firm boundary: whether an expression . . . is to be taken as a genuine proper name . . . depends, not indeed on the impossibility of extruding it from the language without loss of expressive power, but on the extent to which it is embedded in a special vocabulary: this will, of course, be a matter of degree, and we are therefore free to draw the line according to taste (73).

For Dummett, this dangerous position lurks behind his contrast between color words used as nouns and other sorts of abstract nouns derived from adjectives and verbs. This contrast consists partly in the fact that the use of color words

> . . . is linked with that of a special vocabulary of predicates and relational expressions . . . which are used either only in this connection or else in special senses which have to be learned (72).

Dummett avoids this threat by finding more to this contrast than the above-mentioned fact; he appeals to a (rather theory-laden) notion of identification and an associated "well-understood use of demonstratives." How well this response fares for color words is not now at issue; for it's unclear how far the analogy between color words and numerical expressions may be pushed. Dummett admits that

> . . . anything that can be said by means of this special vocabulary, with color-words used as nouns, could be re-expressed by sentences in which the corresponding color-words appeared only as adjectives; in some cases the transformation would be easy, in others it would depend on a thorough understanding of the principles of application of these predicates: but it would in no case consist merely in a conversion of one general idiom into another (73).

Perhaps[18]; but if so, the same does not apply to numerical expressions. The sort of transformation exemplified by (vi) to (v) and (iii) to (ii) is, I think, mere "conversion" of idiom.

So the threat remains. In place of an appeal to an unanalyzed (and in this context, inapplicable) notion of identification, I propose to answer the troublesome relativist by appeal to our need for a holistic assignment of logical forms which meshes with an account of reference. It must be possible to view the language of elementary number theory as an integral part of our total language, which includes diverse kinds of discourse. Under this holistic constraint, units of discourse will seek their appropriate levels. We must distinguish a strict notion of logical form subject to this holistic constraint from a looser localized notion applicable to particular arguments or particular kinds of discourse (e.g., that of the number theorist) in isolation from the rest of language. The purpose of this local notion is to give a structural analysis of the relevant corpus of sentences which illuminates the logical relations among the sentences in that corpus; this task does not require "meshing" with an over-all plausible theory of reference. So for an appropriate corpus of sentences (e.g., that of number-theoretic discourse) a sentence whose logical form (in the strict sense) is '$(PX)((Qx)(Xx))$' may be assigned the form '$P(Q)$'; regarding numerical expressions as of type 0 and predicates like 'is prime' or 'is less than' as of level 1 does articulate all the logical relations among purely number-theoretic propositions. However, in assessing ontological commitments, and in doing what Frege would have called "analyzing the structure of thoughts," we take as our unit of discourse our entire language, including sentences like (ii) and (iii). This is the unit for which we owe a debt to the theory of reference; to pay it we must adhere to the strict notion of logical form; otherwise we introduce a perplexing haze around the relationship between sentences like (ii) and those like (iii) or (vi). So the dogmatic Fregean is more right than wrong in denying that numerical terms can stand for numerical quantifiers, for there cannot be a language in which object-quantifiers and objects are simultaneously viewed as having level zero.

The significance of differences of level is brought out by the question: What is it to have a numerical object-quantifier in one's referential ken? Though many papers have been written on the "puzzle": How can we refer to numbers if we don't causally interact with them? to my knowledge no one has been seriously puzzled

[18] For an argument to the contrary, see Armstrong, *op. cit.*, pp. 58-61.

by the question: How can we refer to the existential quantifier if we don't causally interact with it? Only a Fregean would say that '$(\exists x)$' refers to a quantifier at all. If this locution sounds disturbingly odd to you, that fact only bears witness to the point being made. Numerical object-quantifiers are in the same acausal boat as the familiar existential object-quantifier. It is helpful to retain the Fregean way of speaking; part of any satisfying account of reference would describe those aspects of our linguistic and cognitive practice in virtue of which '$\exists$' and other existential-quantifier expressions stand for the existential quantifier. The microstructure of our access to such quantifiers needs to be understood, though it is importantly different from the microstructure of our access to objects, or even to concepts and relations of level 1. The case of Adam is impossible simply because the similarities between our linguistic practices and his make our quantifier-expressions co-referential with his. So it would be a mistake to construe the impossibility of the case of Adam as implying that reference is irrelevant to mathematical truth; it implies only that the "simple-minded" theory of how reference relates to mathematical truth is not acceptable.

The strict logical form of arithmetic sentences like (iii) and (v) is not what their surface grammar suggests. In making what appears to be a statement about numbers one is really making a statement primarily about cardinality object-quantifiers; what appears to be a first-order theory about objects of a distinctive sort really is an encoding of a fragment of third-order logic. Frege late in life seems to have reached this conclusion. In a diary dated March 23, 1924, he wrote:

> My efforts to become clear about what is meant by number have resulted in failure. We are only too easily misled by language and in this particular case the way in which we are misled is little short of disastrous. The sentences 'Six is an even number', 'Four is a square number', 'Five is a prime number' appear analogous to the sentences 'Sirius is a fixed star', 'Europe is a continent'—sentences whose function is to represent an object as falling under a concept. Thus the words 'six', 'four', and 'five' look like proper names of objects, and 'even number', 'square number', and 'prime number' along with 'number' itself, look like concept-words; so the problem appears to be to work out more clearly the nature of the concepts designated by the word 'number' and to exhibit the objects that, as it seems, are designated by number-words and numerals (Hermes *et al.*, 263).

Then in an entry dated March 24, he continues:

> Indeed, when one has been occupied with these questions for a long time, one comes to suspect that our way of using language is mislead-

> ing, that number-words are not proper names of objects at all and words like 'number', 'square number', and the rest are not concept-words: and that consequently a sentence like 'Four is a square number' simply does not express that an object is subsumed under a concept, and so just cannot be construed like the sentence 'Sirus is a fixed star'. But how then is it to be construed? (Hermes *et al.*, 263/4)

The answer to this deathbed question is, I think, suggested by the conclusion of the quotation of pp. 139/40 above: 'Four is a square number' involves a type 3 expression completed by a type 2 expression; it has the form: (Square X)$((\exists 4x)Xx)$. But, one may ask, Why do we speak a language that is so misleading? Why does the familiar local sort of logical form which the mathematical-object picture suggests work so well? Indeed, how does the mathematical-object picture function in our mathematical thinking and our mathematical practices?

The mathematical-object picture may be described in two equivalent ways. We may see its acceptance as a willingness, in certain contexts, to regard type 2 expressions (and their referents) as if they had type 0. Or we may see it as the pretense of positing objects that intrinsically represent certain type 2 entities. This second description makes mathematical discourse, when carried on within the mathematical-object picture, a special sort of fictional discourse: numbers are fictions "created" with a special purpose, to encode numerical object-quantifiers and thereby enable us to "pull down" a fragment of third-order logic, dressing it in first-order clothing. I'll sketch a systematic reconstruction of this stunt. But first we should ask: Why would such a stunt be useful or natural?

Higher-order logic is notationally messy and logically complex. For purposes of everyday life, and even for advanced research in pure number theory, there is no need to express arithmetic propositions in a notation that exhibits the higher-order nature of the thoughts involved. Such a "coding device" loses nothing (except freedom from philosophical confusion) and gains much. For first-order logic is a many-splendored thing, worthy of the high pitch of Quinean rhapsody:

> Classical quantification theory enjoys an extraordinary combination of depth and simplicity, beauty and utility. It is bright within and bold in its boundaries.[19]

It is familiar—the logic we learn on our parents' knees, and then really learn in our first logic class; it is completely axiomatizable and compact and has an exceptionally tractable model theory. Phil-

[19] *Ontological Relativity* (New York: Columbia, 1969), pp. 112/3.

osophical rigor does not require that we abandon these advantages to first-order mathematical discourse, but only that we see it right.

We now run up against a more general problem: the logical status of fictional discourse and various kinds of discourse "about" fictions. In adopting the slogan "Numbers are fictions," are we referring to numbers and attributing to them fictionality? Or are we attributing to a body of discourse the attribute of being apparently, but not really, referential?

Saul Kripke[20] and John Searle[21] wish to reify fictional characters, at least in construing the discourse of literary historians and critics. On this view, in creating the play *Hamlet,* Shakespeare created a host of other abstract objects, namely the various characters in the play. Kripke has argued that this doctrine is not the superficially similar doctrine frequently attributed to Meinong. Even on this view, there may be contexts in which fictional names are not to be construed as designating; e.g., when used by a storyteller who creates his characters as he invents his tale.

I'll remain neutral on this issue. A coding-fictionalist willing to posit fictional characters might as well posit numbers; (3) would still be rejected. For numbers would be created by certain sorts of intellectual activity: had no one formulated mathematics in terms of such fictions, they would exist only as "permanent possibilities of fictionalizing." Such a coding-fictionalist is still resisting the mathematical-object theory. (If reification of fictions is combined with reification of possible objects, including possible fictions, then coding-fictionalism seems to become a version of the mathematical-object theory; classification of such a position will be left to those who find it attractive.)

There are significant differences between mathematical discourse within the mathematical-object picture and the familiar fictional "histories" of myths, novels, etc. The storyteller who creates as he goes does not make genuine assertions or even express propositions with truth values (for the most part); he therefore has a free hand in a way in which the ur-mathematician, who, we may imagine, first introduces mathematical fictions, does not. From the start the ur-mathematician is beholden to a body of truths, e.g., truths of third-order logic; rather than pretend to make assertions, he makes primary assertions indirectly (in Searle's terminology[22]), by pretending to make secondary assertions "about" fictions.

[20] In a talk given at Cornell in 1982.

[21] "The Logical Status of Fictional Discourse", reprinted in *Expression and Meaning* (New York: Cambridge, 1979).

[22] "Indirect Speech Acts," *ibid.*

Pressing this disanalogy further, it might be urged that the critic or literary historian does (or should) express propositions with truth values; but the propositions expressed will have their truth values contingently, depending on the whims of those who created the fictions under discussion. Mathematical assertions, on the other hand, are not dependent on the whims of the ur-mathematician. But this disanalogy is not so persuasive. Contrast the ur-mathematician with Shakespeare, who one day invents the core plot of *Hamlet*—a tale of a prince who comes to believe that his uncle has murdered his father . . . —, casts around for a name for the protagonist, and hits on 'Hamlet'. Suppose that he in no way draws on history, prior literature, or legend. He creates a plot, and then a play, in which certain things are true. He has thereby made the following literally true:

(vii) In the fiction *Hamlet:* Hamlet came to believe that his uncle murdered his father.

Is this a necessary truth? Could Shakespeare have invented a radically different plot involving that very character? I think not. He could have come up with a different plot and named its protagonist 'Hamlet'; but this is irrelevant.

The following analogy may clarify the special nature of mathematical fictionality. What do I see when I see myself in a mirror? If I answer "A mirror image of myself," what is the status of this mirror image? Is it a massless object located on the other side of the mirror, visible only from certain positions in the room? I don't think that my image is among the objects located behind the mirror; rather I think that talk of mirror images is a sort of fictional discourse. Statements "about" such fictions are not made true or false by our whims; rather they "encode" facts about the things reflected in mirrors. In saying that I'm looking at my mirror image, I say that I'm looking at myself. Could we identify my mirror image with myself? If we're discussing only objects reflected in the mirror, this may be helpful and harmless; but if we're also discussing objects not reflected in the mirror, this identification invites confusion. Similarly, if our discourse includes sentences of the forms '$(Qx)Ax$' and '$(PX)((Qx)Xx)$', we can't regard 'Q' as an expression of type 0. Certain purposes, e.g., understanding the logical structure of pure number-theoretic discourse in splendid isolation, permit us to ignore sentences of the former forms; for such restricted purposes, we can regard 'Q' as a singular term. Within a global account of reference, there is no simple way truthfully to complete the schema:

Numerical terms refer to ---s.

I'll now give a sketch of the model theory for a language that formulates cardinal arithmetic in accordance with the mathematical-object picture. Rather than stand for a particular numberer, 'the number of' in effect contains a hidden free-variable ranging over numberers; a sentence involving 'the number of' is like an open formula, true iff satisfied by all numberers. The arithmetic sentences of "real-life interest" are satisfied by all numberers iff they are satisfied by any—this is why it's harmless, and in fact helpful, to pretend that 'the number of' stands for a particular "standard" numberer.[23] But some sentences, like (iv), are satisfied by some numberers and not by others. This selectivity is the basis for the natural "intuition" that (iv) has no truth value. The pull to say that (iv) is false is a case of our tendency to construe the truth of a formula as satisfaction by any choice of values. (This explains the peculiarity of (iv) in a way in which mere appeal to category mistakes cannot; for why do the alleged category boundaries fall where they are alleged to fall?)

A set A is *acceptable* iff card $\{\text{card}(x) \mid x \subseteq A\} \leq \text{card}(A)$, where card (x) is the Scott cardinal of a set x; equivalently, iff A satisfies (D). Where $\mathfrak{L}$ is a first-order language, let $\mathfrak{L}^{\#}$ be the result of enriching $\mathfrak{L}$ with the predicates 'Number (ν)' and '$\nu \leq \mu$' and the term-forming operator '$(\#\nu)$'; '0' and '$n + 1$' will abbreviate '$(\#\nu)(\nu \neq \nu)$' and '$(\#\nu)(\nu \leq n)$'. Where $\mathfrak{A}$ is a model for $\mathfrak{L}$, $\mathfrak{N}$ is an $\mathfrak{A}$-representor iff $\mathfrak{N}$ is a one-one function from $\{\text{card}(x) \mid x \subseteq |\mathfrak{A}|\}$ into $|\mathfrak{A}|$; satisfaction and denotation for formulas and terms of $\mathfrak{L}^{\#}$ are defined relative to $(\mathfrak{A}, \mathfrak{N})$; satisfaction in $\mathfrak{A}$ is satisfaction in all $(\mathfrak{A}, \mathfrak{N})$ where $\mathfrak{N}$ is an $\mathfrak{A}$-representor; implication and validity are now defined as usual; a precise explication of thesis (1) is under way.[24]

[23] Discussions of the United States Constitution may contain occurrences of 'The President' which only a rabid Meinongian would construe as designative. Many mathematicians will speak of "the countable atomless Boolean Algebra" without intending to refer to a particular structure, relying on the fact that all such structures are isomorphic. Such phrases may also be construed as involving a free variable with restricted range. I think that talk of types and tokens is best understood along these lines. Armstrong (xiii, 16/7) conflates the type-token distinction with that between universals and particulars. Unlike such pseudo-definite descriptions as 'The property of whiteness', which at least have the decency to be obvious nominalizations of predicates, typical terms that designate types, e.g. 'The 1948 D penny', 'The second edition of *Principia Mathematica*', 'The letter 'A'', etc., are primarily singular terms—a point which should not be ignored in describing their semantics. Arithmetic discourse is more misleading than talk of the countable atomless Boolean Algebra or the 1948 D penny only because numberers tend to be occluded by the numbers themselves; once numberers and representors have been brought into the foreground, the analogy should be clear.

[24] This is all said "within" the mathematical-object picture; I have discussed only discourse "about" cardinal numbers, not "about" sets; that will be discussed elsewhere. Note: the Axiom of Choice implies that all infinite sets are acceptable.

III. TWO PROBLEMS

The position I have defended in the previous section is rather like that of Russell and Whitehead. They too thought arithmetic to be higher-order logic; they viewed talk of sets as abbreviatory for talk of propositional functions.[25] And the two accounts share two difficulties.

We've assumed that cardinal numbers number only the objects falling under type 1 concepts. But need this be the case? Provided that we replace identity by co-extensionality of the appropriate type, it seems that "there are four concepts under which at most Tom and Jane fall" is intelligible and, in fact, true. More generally, we seem to have numerical quantifiers over every type.

Russell and Whitehead, who come close to identifying numbers with numerical quantifiers, concluded that cardinal numbers were typically ambiguous: for each type σ there are numbers of type $((\sigma))$. David Bostock maintains (incorrectly, I think) that there are "pure" quantifiers, outside of the Fregean type structure, and that numbers are such quantifiers.[26] My lesson is different: we should recognize representors defined on numerical quantifiers of each type $((\sigma))$, with values of type 0, paired with numberers defined on concepts of type (σ); '#' is typically ambiguous, disambiguated by the type of the variable it binds; '$(\#\nu^{\sigma})\phi$' is of type 0. Of course a system of representors and numberers must "mesh" together properly: where '$(Q_E\nu^{\sigma}\nu^{\tau})$' represents equinumerosity between types σ and τ, we must have:

$$\text{(A}'\text{)} \qquad (\#\nu^{\sigma})\phi = (\#\nu^{\tau})\psi \equiv (Q_E\nu^{\sigma}\nu^{\tau})(\phi(\nu^{\sigma}),\ \psi(\nu^{\tau}))$$

Our second problem is more serious. Like Russell and Whitehead, we have been assuming that there are infinitely many objects. By abandoning the mathematical-object theory, we lose all mathematical justification for this assumption. Appeal to physics, e.g., to the supposed infinity of space-time points, is rather flimsy, especially if we think that space-time points are also questionable entities. Arithmetic should be able to face boldly the dreadful chance that in the actual world there are only finitely many objects. If this were the case, there would be no numberers: (D) is not satisfiable in a finite domain. But worse is in store. For suppose there were only eleven objects. Then '$5 + 7 = 11$', which is to say:

$$(\forall X)(\forall Y)(((\exists 5x)Xx \,\&\, (\exists 7x)Yx \,\&\, \sim(\exists x)(Xx \,\&\, Yx)) \supset (\exists 11x)(Xx \vee Yx))$$

[25] At least in the second edition of *Principia Mathematica.*

[26] In *Logic and Arithmetic,* vol. I (New York: Oxford, 1974).

would be true. Although such news may gladden the hearts of those who believe mathematical knowledge to be in principle revisable, we're still committed to the truth of (vi), that is, of '$5 + 7 = 12$'; and this will not do.

For a long time it has been thought that modality played no role in mathematics, since purely mathematical truths were uniformly necessary. The currency of this impression testifies to the power of the mathematical-object picture, which enables us to ignore modality by taking up the slack into its bloated ontology.[27] But modality really permeates the terms in which we learn and discuss mathematics. For example, if we ask a bright child what it means to say that there are infinitely many numbers, the answer we want is something like "No matter how high I were to count, I could go on and count higher."

The answer to our difficulty is as old as Aristotle: The notion of infinity required by mathematics is merely that of a potential infinity. Even if at every possible world only finitely many objects exist, still any world has access to a richer world with some more objects. Assuming that our underlying modal logic includes S4, only the slight addition of a '$\Box$' is required to readjust our analysis of simple arithmetic statements. For example, we replace (vi) by:

$$(\forall X)(\forall Y)\,\Box\,(((\exists 5x)Xx \,\&\, (\exists 7x)\,Yx \,\&\, \sim(\exists x)(Xx \,\&\, Yx)) \supset (\exists 12x)(Xx \vee Yx))$$

Note: Variables of type 1 now range over attributes which at any possible world collapse to Fregean concepts. We have deftly dodged the danger posed by sparsely populated worlds. The revision of (1) which I am defending is:

(1′) Mathematics is higher-order modal logic.

Here we have the start of a long story; so let's make this the end of another.

HAROLD T. HODES

Cornell University

[27] See Putnam, "Mathematics without Foundations," reprinted in *Mathematics, Matter and Method* (New York: Cambridge, 1975).