

CAN OUR QUANTIFIERS RANGE OVER ALL COLLECTIONS?

by

THOMAS ANTOGNINI

B.A., Long Island University
(1973)

SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE
DEGREE OF

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

APRIL, 1981

Signature of Author.....

Department of Linguistics and Philosophy, April 30, 1981

Certified by.....

Thesis Supervisor

Accepted by.....

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Chairman, Department Committee

MAR 1 1984 ARCHIVES

LIBRARIES

CAN OUR QUANTIFIERS RANGE OVER ALL COLLECTIONS?

by

THOMAS ANTOGNINI

Submitted to the Department of Linguistics and Philosophy on April 30, 1981 in partial fulfillment of the requirements for the degree
 Doctor of Philosophy

ABSTRACT

Charles Parsons has argued, roughly as follows, that we cannot succeed in quantifying over all collections. We seem bound to adopt a metatheory for ZF in which truth for ZF is definable. But if we try to define truth for ZF without recourse to semantical notions, we must admit quantifiers over proper classes. But then the domain of these class quantifiers may not seem to contain all collections, since we must in turn obtain a truth theory for this class theory. So we seem never to arrive at a theory in which we can talk about all collections.

I contend that we can do quite well with a theory in which satisfaction for the formulas of ZF is inductively defined; this theory will not embroil us in proper classes. We can determine in a sense directly, I argue, that the SAT predicate is well defined.

Parsons has constructed a translation from NB (Von Neumann-Bernays set theory) into ZF plus its truth theory.

Can this translation be taken to provide an ontological reduction, or can it be so modified that it does? If it can, then in adopting ZF plus its truth theory we are perhaps already admitting, however unwittingly, the proper classes of NB. For if we reduce one ontology to another, we in effect show that in adopting the second ontology we are thereby committed to the first.

I propose some constraints on the way in which the relevant structure of a theory must be preserved in that of another for there to be a genuine reduction. On this score, the would-be reduction of NB to ZF plus its truth theory is seen to fall short.

ACKNOWLEDGEMENTS

This thesis owes its existence to a variety of people. On its behalf, I thank Professor George Boolos, who was generous in both insight and encouragement; likewise I thank Professors Richard Cartwright, Sylvain Bromberger, and Hilary Putnam for gracious and penetrating criticism along the way. Helen Shuman deserves praise for marvelously transforming my manuscript into these finely typed pages. Finally, to my wife Francesca I owe a special debt; her unflagging support was remarkable, if no longer surprising.

Chapter I

Paradox and Second Order Quantification

In a number of papers,¹ published several years ago, but unfortunately little discussed in philosophical journals, Charles Parsons presents a case for the view that we cannot succeed in making our quantifiers range over all collections. Much of the evidence for this case consists of considerations arising from the liar paradox. Currently, most of the industry generated by the liar paradox seems to be directed towards Kripke's theory of truth;² this might create the impression that Kripke's theory somehow superceded Parsons' reflections. Such an impression, I am convinced, would be mistaken. For while most recent discussions of the Liar paradox are like Kripke's, concerned with its consequences for the theory of natural language, there is a wider question of its impact on languages in general. Even if we can construe the truth predicate in natural languages so that, given the expressive power of natural languages, no liar paradox can arise, in a formal language with greater expressive power the same technique may not suffice. Insofar as we wish to maintain that such stronger formal languages are meaningful, we are still faced with the problem of avoiding paradox. At this juncture, Parsons' observations become once more relevant.

One way to see precisely how the difficulty described in the paragraph above can arise is to consider the inadequate expressive power of the languages Kripke constructs in his theory

of truth. Hence I will begin this paper by focussing on Kripke's theory. Once the nature of the problem I have indicated is more clear, I will discuss in the body of the paper some aspects of the impact of the liar paradox on ontology, epistemology, and the philosophy of mathematics. In particular, I will address at length the question: must our quantifiers take as values only collections from a limited portion of the mathematical universe? This will involve me in a careful examination of Parsons' arguments on this score. I shall argue that considerations originating from the liar paradox alone should not compel a platonist to the view that quantifiers may not range over all collections.

Kripke's theory of truth is basically an attempt to construe the truth predicate of natural language in such a way as to account for uses of certain English sentences in which 'true' occurs, namely those sentences which talk about other sentences in which 'true' also occurs. Tarski had defined 'true' in his systems in such a way that this sort of thing could not happen.³ On Tarski's theory, if a sentence A talks about the truth of sentence B, then the notion of truth used in sentence A has a higher index than any that is used in B. But then sentence B could not in turn talk about the truth of sentence A, since any notion of truth in B must, in consequence, have both a lower and a higher index than any in A. These restrictions hold for Tarski's systems essentially because in these systems the truth predicate is 'bivalent'; that is, it is always defined, so that for any object in the domain, that object satisfies either the truth

predicate, or its negation. Given that these systems can express their own syntax, the truth predicates must be indexed in a hierarchy in order to escape the liar paradox. Kripke's theory develops suggestions advanced by a number of philosophers (e.g., Martin, Putnam, Harman) that a natural language be considered to contain its own truth predicate, and that paradox be scotched by allowing truth value gaps in the interpretation of the truth predicate. There are various considerations that make it desirable to construe natural language in this way. First, it is quite difficult to understand the claim that natural languages are inconsistent, which is what Tarski believes. In fact, only theories can be inconsistent, and natural languages are not theories. It seems that natural language is best taken as a syntactical system with different interpretations on different occasions. Second, it is evident that natural languages do not possess diverse truth predicates arranged in a hierarchy as Tarski's systems would have it, since there is just the word 'true' (or whatever) and no 'true₁', 'true₂', etc., as in Tarski's languages. Third, there is the sort of phenomenon which Kripke adduced, viz., talk of the kind which went on during Watergate (to be more fully described shortly). Fourth, truth value gaps evidently appear in other contexts in natural languages, e.g., failure of presupposition, vague terms.

Kripke's theory of truth is a precise formulation of a language which has its own truth predicate and 'avoids paradox'

(i.e., has an interpretation). A Kripke language may be constructed in the following way.* One starts out with an interpretation of all the sentences in the language except those with the term 'true' in them. Then, if a sentence is true in this interpretation, its Godel number goes into the extension of the truth predicate; if it is false, its Godel number goes into the antiextension of the truth predicate, where the antiextension of a predicate is the extension of its negation. Once this is done, new sentences are thereby determined to be true or false in the interpreted language, since some sentences in which the term 'true' occurs now will have a determinate truth value. The Godel numbers of these sentences are in turn thrown into the extension and antiextension, respectively, of the truth predicate. This process is iterated as often as possible. Since there are only countably many sentences, it can go on only countably many times before no new sentences are determined to be true or false.

As Kripke points out, in such a language it would be possible to discuss, e.g., Watergate. For in Watergate it happened often enough that a sentence was used, containing the term 'true', which talked about a sentence also containing 'true'. Kripke's theory shows how 'true' in these sentences can be understood univocally. Because there are such situations as Watergate, there is some evidence that 'true' is

*This will be a so-called minimal fixed point Language.

sometimes employed in the manner Kripke indicates. Watergate was not unique in this regard, of course; Kripke's theory could be motivated without Richard Nixon.

Kripke notes that his languages are incomplete in certain ways, however. There are predicates we would want to hold as intuitively meaningful, which cannot be expressed in his language. Now, the fact that there will be some predicates which cannot be captured on a fixed interpretation of a countable language may appear to have a simpler explanation. For there are, e.g., more than countably many subsets of ω ; corresponding to each such subset is a property of natural numbers. But for every property of natural numbers, there is a predicate in some language which expresses that property, that is, has that subset as its extension. Against this, however, observe that while it may be true that there is some such predicate in some language, this language and this predicate may be recalcitrant to any effort to understand them, in that we may have no intuitive grasp of the primitive predicates of this language. It is, of course, difficult to exhibit such a predicate; for once one has fully described a predicate, it must surely fall in among those predicates that we do grasp. But the consideration of cardinality seems to effectively rule out the possibility that we can understand every predicate over the natural numbers. In any event, we can understand the predicates that I take to be inexpressible in natural language.

Thus, for example, some sentences will never get a truth value defined in the process I have delineated, e.g., the liar

sentence "I am not true." We grasp perfectly well what it means for a sentence not to be defined: Kripke himself has explained it; yet in Kripke's languages there is apparently no predicate which has as its extension all and only the sentences whose truth value is undefined. Moreover, there is certainly no predicate in his languages which has as its extension just those sentences which are either false or undefined, and this predicate is also meaningful to us. Hence, insofar as we insist on construing natural language so that it is interpreted in the way Kripke delineates, natural language appears unable to express all intuitively meaningful predicates.

One might try to stave off the conclusion that natural language cannot express all such meaningful predicates. As I have already indicated, it is reasonable to take a natural language to be flexible enough for its predicates to have a variety of extensions in a variety of circumstances. That the analogous thing holds for demonstratives is obvious. One might hope that this flexibility would allow natural language to express each of these meaningful predicates, although of course on distinct uses. The truth predicate in particular, as well as 'x is not true', may on diverse occasions have diverse uses. So one might try to use 'x is not true' at times so that its extension contains precisely those sentences either false or undefined in the Kripke language. Paradox would be warded off because, although the sentence 'this sentence is not true' would be interpreted as being not true₁ (where not

true_1 is the new interpretation of 'not true') no argument shows that the sentence is also true_1 . For one could not infer that since this liar sentence says that it is not true_1 , it therefore says something true_1 , that is, true as on the old interpretation. What it does say is something true_2 , where true_2 is the notion of truth for the language with 'true' interpreted as true_1 .

Now can we proceed to interpret 'true' as true_2 ? I believe that to do this would violate some intuitions about how the liar sentence should be understood. For it would turn out that the liar sentence would not, in an important sense, talk about itself. Thus, suppose we do use 'true' to mean true_2 . As I have remarked, the liar sentence then gets into the extension of 'true' in the sense of ' true_2 ' because it is in the extension of 'not true' in the sense of 'not true_1 '. Hence, the fact that it is not in the extension of 'not true' in the sense of 'not true_2 ' (however the sense of 'not true_2 ' is precisely explicated) is also because it is in the extension of 'not true' in the sense of 'not true_1 '. That is, the liar sentence fails to get into the extension of its own predicate not because of the interpretation its predicate presently has, but instead because of the extension its predicate used to have. Thus, the liar sentence does not talk about itself being under its actual interpretation. Rather, it talks about itself being under a different interpretation. In general, in taking 'not true' to mean not true_2 , any

sentence in which 'not true' appears, and which refers to other sentences in which 'not true' occurs, attributes to these sentences a different interpretation of 'not true' from that which it possesses itself. This situation is unacceptable, inasmuch as what the liar sentence or a similar sentence does in natural language, if it does anything, is to talk about its truth value under the interpretation it in fact has. The only way in which a natural language could accommodate the above sort of interpretation would be if it contained two distinct words, 'true' and 'true₁', and surely this no natural language does. It may seem odd that the expressive limitation of natural language should arise out of what appears to be a syntactical limitation, the absence of 'true₁', or the like, from the vocabulary. But given that 'true' cannot accommodate every notion of truth, as I have just argued it cannot, it is just such a "syntactical" limitation that would confine the expressive capacity of natural language. Observe, moreover, that adding only the words 'true₁' and/or 'true₂' to English would not really ameliorate the situation. The sort of problems which led to the introduction of 'true₁' and/or 'true₂' would visit in turn this new language, and call for the annexation of 'true₃', 'true₄', etc.

At any rate, that not all intuitively meaningful predicates can be expressed in a natural language is not, as far as I can see, any cause to despair. There is no apriori reason to surmise that it should be otherwise. In any event, there plainly

are formal languages in which these diverse predicates can be expressed. But the issue then arises; is there any problem dealing with the liar paradox within formal languages? It would seem, in the face of Tarski's way of defining truth, that there is not: for apparently, we can always set up truth theories for a language, a la Tarski, in which it is understood that the truth predicate does not apply to the meta-language itself. This view at first glance seems to offer comfort and security, but look again. In the case of set theory, it is not clear what a Tarski-style truth theory comes to. Tarski showed how, in languages of a restricted class; truth theories can be constructed. If the range of the quantifiers is a set, and we understand the primitive predicates of a language, then we can set up a truth theory for that language, without recourse to semantical notions. However, on a standard construal of the language of ZF, the range of the quantifiers is not a set, not a collection of any kind. There will be no way of eliminating semantical notions in a truth theory in the typical Tarski fashion. An example of a typical Tarski-style truth theory may help to explicate this point. Suppose a truth theory for the language of arithmetic is sought. The range of the quantifiers is a set, the set of all numbers. The Tarski style truth theory is effected in a theory in which some set theory is present. First, a conjunction of open formulas, call it $A(x)$, is defined:

- 1) $\forall n \forall s \forall i \forall j \forall k (n = \ulcorner x_i + x_j = x_k \urcorner \rightarrow (\langle n, s \rangle \in X \leftrightarrow (s)_i + (s)_j = (s)_k))$
- 2) $\forall n \forall s \forall i \forall j (n = \ulcorner (x)_i' = x_j \urcorner \rightarrow (\langle n, s \rangle \in X \leftrightarrow (s)_i' = (s)_j))$
- 3) $\forall n \forall s \forall i (n = \ulcorner x_i = 0 \urcorner \rightarrow (\langle n, s \rangle \in X \leftrightarrow (s)_i = 0))$
- 4) $\forall n \forall s (n = \ulcorner \psi \vee \phi \urcorner \rightarrow (\langle n, s \rangle \in X \leftrightarrow (\langle \ulcorner \psi \urcorner, s \rangle \in X \vee \langle \ulcorner \phi \urcorner, s \rangle \in X)))$
- 4) $\forall n \forall s (n = \ulcorner \neg \phi \urcorner \rightarrow (\langle n, s \rangle \in X \leftrightarrow \langle \phi, s \rangle \notin X))$
- 6) $\forall n \forall s \forall i (n = \ulcorner \exists x_i \phi \urcorner \rightarrow (\langle n, s \rangle \in X \leftrightarrow \exists s' [s' = s \text{ except, possibly, at the } c^{\text{th}} \text{ place} \& \langle \ulcorner \phi \urcorner, s' \rangle \in X])$
- 7) $\forall n (n \in X \rightarrow n \text{ is the Godel \# of a formula}) \quad \& \langle \ulcorner \phi \urcorner, s \rangle \in X$

An explicit definition of satisfaction can be given:

$$\text{Sat}(x, s) \leftrightarrow \exists! y [A(y) \& \langle x, s \rangle \in y]$$

From this, we get truth for closed formulas:

$$\ulcorner \phi \urcorner \text{ is true} \leftrightarrow \exists s \text{Sat}(\ulcorner \phi \urcorner, s)$$

Why does this method miscarry for the language of ZF?

Well, let us endeavor to succeed. Here is the obvious attempt:

Define A(x) as the conjunction of (1) - (4) below.

- 1) $\forall s \forall n \forall i \forall j (n = \ulcorner x_i \in x_j \urcorner \rightarrow (\langle n, s \rangle \in X \leftrightarrow (s)_i \in (s)_j))$
- 2) $\forall s \forall n (n = \ulcorner \psi_1 \vee \psi_2 \urcorner \rightarrow (\langle n, s \rangle \in X \leftrightarrow (\langle \ulcorner \psi_1 \urcorner, s \rangle \in X \vee \langle \ulcorner \psi_2 \urcorner, s \rangle \in X)))$
- 3) $\forall s \forall n (n = \ulcorner \neg \psi \urcorner \rightarrow (\langle n, s \rangle \in X \leftrightarrow (\langle \ulcorner \psi \urcorner, s \rangle \notin X)))$
- 4) $\forall s \forall n \forall i (n = \ulcorner \exists x_i \psi \urcorner \rightarrow (\langle n, s \rangle \in X \leftrightarrow \exists s' (s' = s \text{ except, possibly, at the } c^{\text{th}} \text{ place} \& \langle \ulcorner \psi \urcorner, s' \rangle \in X)))$

This condition A(x) is true in ZF of no set whatsoever; hence

$\exists! y (A(y) \& z \in y)$ is satisfied by no set. A(x) has no solution

because for condition (1) to be satisfied by a set, that set would have to contain pairs of formulas and sequences where the sequences are of arbitrarily high rank. This no set can do.

Hence, we cannot hope to do a strict Tarski-style truth theory in which we can do away with semantical notions, but must settle for some rough analogy. What this analogy is, we shall

presently consider. Let us remark in advance, however, that certain hard problems lie in wait. Are there theories which we are willing to accept, but whose truth theories involve principles that we cannot justify, and whose truth theories we therefore are indisposed to admit? If we do assent to something like a Tarski style truth theory for any theory, does this decision have any undesirable ontological consequences? Parsons has shown that what he calls NB (Von Neumann Bernays set theory) and even NB+ (the extension of NB that allows impredicative class formulas in set separation axioms) are translatable into ZF plus a certain Tarski-like truth theory for ZF. The ontology outstrips that of ZF; must we concede to this ontology if we accept ZF plus its own truth theory? Parsons at times is quite sympathetic with the view that we should; he proposes at one point that assenting to such an ontology might require in turn admitting an even richer ontology, and further that the ontologies we should embrace are indeed essentially open ended. Parsons even goes so far as to proffer this as possible evidence that the quantifiers of set theory should be understood in a quasi-intuitionistic manner.

Still another difficulty is this. Assume there is an openended character to the theories we are willing to affirm, viz., when we hold a theory we will hold its Tarski-like truth theory. How should discourse about all theories, or all interpretations be construed, when ostensibly such discourse must be in the language of some particular theory? Paradox

appears to be close at hand. This problem is in a way the most intriguing of all, but it owes its special intrigue to its being the most perplexing. Fortunately for me, it is not germane to the central issues of this paper.

Now, before attending to the question, 'What kind of truth theory, similar to that of Tarski, can be constructed for set theory?' it is important to see that a truth gap truth theory is of no avail in any attempt to evade the problems facing the construction of a Tarski like truth theory. For there would be no point insisting on a Tarski like truth theory for set theory if a less problematic alternative exists. One way to see the futility of employing a truth-value gap approach to handle these problems is to answer the question: is there a way out of the liar paradox, that works for both natural and formal languages, by means of some truth value gap approach? I shall show that, in an important sense, there cannot be, although of course there doubtless is some resolution of it for natural languages (perhaps Kripke's theory constitutes such a resolution). What I wish to argue is that with very minor assumptions about what predicates we find meaningful, a language under a fixed interpretation cannot express all of the meaningful predicates there are. This is the sense in which the liar paradox cannot be got around. One might put this conclusion thus: there are no universal languages.

Suppose: (1) if we find a language meaningful then we

can create a predicate and a language and interpret them in such a way that the extension of the predicate in that language will be exactly the true sentences of that first language; (2) if there is a language in which a predicate is meaningful and has a certain extension, then there is a language in which that predicate occurs, with the same extension, and in which there is what I shall call the complementary predicate of that predicate. The complementary predicate A_1 of a predicate A_0 is a predicate which has in its extension just those things which A_0 does not. In such a metalanguage it is impossible that the predicate having as its extension the true sentences in the object language could be expressing its own truth predicate. Suppose otherwise. Then the sentence 'This sentence is not true' (where 'not true' is the complementary predicate of 'true') could not be in the extension of the truth predicate or of its complementary predicate, yet by hypothesis must be in at least one. The metalanguage, hence, must have greater expressive power than the object language.

It is noteworthy that this result holds in a very wide variety of cases. Thus, even if there were more than one way for a predicate to be undefined (even infinitely many), the possibility of such a language being its own metalanguage is precluded, for no mention is made of undefined predicates in the statement of the conditions.

How plausible are these conditions? Certainly prima

facie they are eminently reasonable. Consider condition (1). It is difficult to see how we could find a language meaningful, and yet not grasp the notion: sentence which is true in the language. But it seems that our grasping this notion consists in, or at minimum would involve, being able to intend some language so that one of its predicates has as its extension just the true sentences of the first language.

Condition (2) also seems ineluctable. If we can understand a predicate, we are able, it would appear, to intend the predicate in a determinate way to apply to certain entities and not to apply to others. But this is in some sense to split the universe into two mutually exhaustive parts. And if we can envision the universe as being thus split, then surely we could employ a language in which there are two predicates, one of which has as its extension one part of the split, and the other, the other part. But this is just condition (2). In fact, our way of construing predicates seems to be closed in other ways, too. Thus, there is closure under unions and intersections as well as complementations.⁴

Granted, then, the inevitability of these two conditions, we are driven to conclude that the series of languages that we should acknowledge as meaningful is essentially openended. This, I claim, entails that the series of truth theories we should assent to is also openended. For if a metalanguage for a language is meaningful to us, there is reason to surmise

that that is so in virtue of our having adopted the corresponding truth theory for that language. Or, more weakly, if we hold the metalanguage for an object language to be meaningful, at the very least we appear to be obliged to hold the corresponding truth theory for that object language. Since, as I have argued, there is an openended series of metalanguages we find meaningful, there must be a corresponding openended series of truth theories that we are obliged to adopt.

The upshot of the above is this. Given any theory, we are willing to adopt a certain metalanguage for the language of that theory. In this metalanguage, there is a bivalent truth predicate strong enough to provide a notion of truth for the language of the theory. This bivalent truth predicate is not captured in the Kripke language. Moreover, the truth theory for this theory will be in the style of Tarski, at least in the important respects, and we are obliged to accept that truth theory. Hence, a truth theory analogous to Tarski's is de rigueur, however troublesome to formulate in particular cases, as with the case of set theory.

There is a more fundamental reason that Kripke's theory of truth offers no salvation from the difficulties with setting up a typical Tarski truth theory for the language of set theory. The first step in the process by which the truth predicate starts to get its partial definition a la Kripke is this. The sentences in the ground language are

determined to be true or false by the interpretation that is given them; then the Gödel numbers of these sentences are placed in the extension and anti-extension, respectively, of the truth predicate. But if the ground language under question is the language of set theory, from the standpoint of what background theory are we going to determine that any given sentence of set theory is true? To do this is, in effect, to have a Tarski like bivalent truth theory at hand for the language of set theory; but it is precisely the problem of getting this that now vexes us.

However, my points regarding the inadequacy of a truth value gap approach are still relevant. For suppose we could somehow get our hands on a Kripke language for set theory, and we were satisfied that all the notions of truth we would ever desire to express were expressed in that language. Suppose also that we grant Parsons' suggestion that if we adopt a Tarski like truth theory for the language of a theory, we must admit subcollections of the domain of the quantifiers in that language. Then, although the quantifiers in the language of ZF might not be able to range over all collections, it is at least conceivable that there would be only α many more ranks of collections, where α is the fixed point for the Kripke language. Thus, corresponding to each level in the construction of a Kripke language would be a Tarski-style truth theory for the language of that level; and corresponding to each such truth theory there would be a new rank of

collections to which we commit ourselves, as Parsons recommends. When the levels run out, so do the ranks. Perhaps some new set theory, with distinct quantifiers for each of these α many new levels, would have quantifiers ranging over all collections. But, of course, all this would be a possibility only if the suppositions with which I started were true, and at least the first is not.

Let us resign ourselves then to an openendedness in the truth theories we are willing to accept. Now at first blush this openendedness might be thought to be purely ideological. For in accepting a stronger theory we are not in general compelled to embrace a larger ontology. However, Parsons has set forth some reasons (which I have sketched) to hold that, in the case of truth theories for set theories, we should embrace stronger ontologies as we embrace stronger ideologies. I shall now examine these reasons.

As I have already observed, Parsons shows that a truth theory for ZF plus ZF is of the same strength as the extension of von Neuman-Bernays set theory by allowing impredicative class formulas into the set separation axioms; the two theories are intertranslatable. Parsons considers this to be evidence that accepting the truth theory for ZF involves us in adopting also the ontology of NB+. However, why this should be so is not clear. It is certainly not obvious that the existence of the translation shows that we must adopt NB+ or even NB if we adopt ZF plus its truth

theory. It is true that even in accepting the truth theory for ZF plus ZF we will be constrained to embrace certain entities which were not, in a certain sense, forced on us by ZF alone. Thus, in ZF plus its truth theory, separation is strengthened by allowing formulas composed out of the satisfaction predicate for ZF to separate off sets. This strengthening must be allowed if we are to prove certain trivial facts about ZF, e.g., that all provable formulas of ZF are true. So we can prove in this theory (which henceforth will be called ZFT) that certain sets exist which in ZF alone we were unable to prove. However, the sense in which we must accept new entities in this case is apparently quite different from the sense in which we must if we adopt NB or NB+. For we consider the quantifiers of ZF to range over all the entities which we can prove to exist in ZFT; we are merely unable to prove in ZF that these entities exist. In contrast, the second order quantifiers of NB or NB+ range over entities that seemingly could not be in the range of the quantifiers of ZF, or ZFT.

Granted this difference between the theories NB and NB+ on the one hand, and ZFT on the other, what are the considerations which would lead us to adopt ZFT or NB or NB+?

To appreciate the significance of this question, perhaps it is best to consider the consequences of the view that since we should adopt ZFT, we should admit also NB or NB+.

On this view, we do not by adopting ZF alone succeed in having our quantifiers range over all collections; the proper classes of NB and NB+ elude our attempt to quantify over all collections. And a further consideration makes our predicament particularly vexing. It would appear that the same reasons which led us to adopt NB or NB+ will lead us ultimately to adopt a larger ontology than that of NB(+). For insofar as the truth theory for NB(+) is a theory we should accept, there is again the issue: should we accept the extension of NB(+) by collections over all the classes of NB(+), or NB(+) plus the truth theory for NB(+)? Evidently, if there was reason to opt for NB(+) over ZFT, this reason will suffice to motivate the higher order extension of NB(+) just described as the appropriate choice. Since the truth theories which we should adopt are openended, and since in the train of each new truth theory would come an expanded ontology not in the range of the quantifiers of the previous theories, the range of our quantifiers seems to be openended. It is this situation which leads Parsons to suggest that the quantifiers of set theory might best be understood in a quasi-intuitionistic manner.

It behooves us, then, to see how cogent are the reasons for choosing NB or NB+ over ZFT at the very onset. I shall argue that for one with strong platonist leanings, there is no compelling consideration to prompt the adoption of NB or NB+, or rather that there is none arising out of the adoption of ZFT.

The platonist usually believes that we can use our quantifiers to take on any collection as a value, if we so intend our quantifiers. For the domain of all collections is a well determined totality which exists independently of us, and insofar as we can succeed in referring to mathematical objects at all, we would seem to be able to talk about everything in this totality at once. There are those philosophers, for example Jonathan Lear⁵ and Parsons,⁶ who assert that we cannot quantify over all collections because we cannot have certain kinds of intentions towards all collections, and failing these intentions toward a collection we cannot have it as a value of our variables. Against such a view I will argue at a later time. I will now deal with the question whether a standard, staunch platonist who wants to maintain that we can make our variables range over all collections should be moved by considerations coming from truth theories to give up this position. I think not; let me explain why.

I shall begin by showing in some detail what ZFT is like. To start with, one adds to the language of ZF a dyadic predicate 'Sat(x,y)'. One then defines satisfaction in ZF inductively thus:

- 1) $\forall s \forall n \forall i \forall j (n = \ulcorner x_i \in x_j \urcorner \rightarrow (\text{Sat}(n, s) \leftrightarrow (s)_i \in (s)_j))$
- 2) $\forall s \forall n (n = \ulcorner \neg \psi \urcorner \rightarrow (\text{Sat}(n, s) \leftrightarrow \neg \text{Sat}(\ulcorner \psi \urcorner, s)))$
- 3) $\forall s \forall n (n = \ulcorner \psi_1 \vee \psi_2 \urcorner \rightarrow (\text{Sat}(n, s) \leftrightarrow (\text{Sat}(\ulcorner \psi_1 \urcorner, s) \vee \text{Sat}(\ulcorner \psi_2 \urcorner, s))))$
- 4) $\forall s \forall n \forall i (n = \ulcorner \exists x_i \psi \urcorner \rightarrow (\text{Sat}(n, s) \leftrightarrow \exists s' = s \text{ except, possibly, at the } i^{\text{th}} \text{ place \& Sat}(\ulcorner \psi \urcorner, s')))$
- 5) $\forall n \forall s (\text{Sat}(n, s) \rightarrow n \text{ is the Godel \# of a formula of ZF})$

Inasmuch as we are already committed to ZF, we can consider the clauses of this inductive definition of satisfaction to be added as axioms to ZF. As usual, truth is defined for any closed formula ϕ as follows: ϕ is true if there exists a sequence s such that $\text{Sat}(s, \phi)$. However, we should note that this system just as it stands is inadequate to prove certain elementary facts about ZF: e.g., that all the theorems of ZF are true. To prove these facts, it is necessary, as I have previously noted, to allow formulas in which the Sat predicate appears to be used in the axiom schema of separation. Now is there any way of justifying this further extension of ZF? Moreover, in order to show even that all the Tarski biconditionals are provable, we must permit use of mathematical induction with formulas built up out of the Sat predicate: how is this justified? Parsons finds this second question difficult. As it turns out, there is an important connection between the two questions.

Let us deal with the first question; this will lead naturally into a discussion of the second. I am persuaded that, if we allow the Sat predicate in separation, it is in virtue of a certain kind of mathematical induction involving the Sat predicate that we do so. For, as I shall argue, there is a use of mathematical induction that justifies the belief that Sat is a well defined predicate; and a predicate may be employed in separation just in case that predicate is well defined. This last claim warrants some explanation.

In order to permit a predicate in separation, it would seem enough that the predicate is used so that it applied determinately to each object in its domain. That is, we must assure that its usage is not such that there is an object of which the predicate cannot either be said to apply or said not to apply. That this condition should suffice comports well with our intuitions about sets. In ZF, once a set appears at a certain rank, all of the subsets of that set appear also, and any formula of set theory, even one in which there are parameters and unbounded quantifiers, can separate off a subset of that original set; or, more precisely, such a formula will allow us to see that this subset must exist. Because of what Paul Bernays has called the combinatorial nature of sets (whereof more later), we can see a subset of a given set must exist, so long as we can conceive some series of "decisions", putting members in and excluding members from the subset. This series must, of course, be consistent and apply in a determinate way to each object in the original set. There are predicates of ZF that contain unbounded quantifiers for which nothing short of checking the entire universe of sets will allow us to determine whether they hold of a particular set. Nonetheless, since the formula is used in such a way that for every value of the free variable it is fixed whether it holds or not, we can take that formula to provide decisions about each member of the starting set whether or not it is to be in

the subset; hence the subset should be said to exist.

Now there is nothing special in all this about the formula that performs the separation being a formula of set theory. Any formula that made these decisions in a similar fashion would cut off subsets also. In particular, then, if the Sat predicate is thus well determined in its application, it too should be allowed in separation. Some predicates do occasion doubt as to how satisfactory they would be in separation, because we lack confidence that their usage is determinate. Later in this paper I will discuss in some detail a certain predicate 'R(x,y)' that is to be true of each nonempty set and precisely one member of that set. There I argue that the predicate is not so definite in its usage that we can really be said to pick out a unique interpretation for it. For this reason, 'R(x,y)' would seem unsuitable as a predicate to be used in separation (in fact, it seems unsuitable as a predicate to be used at all, as I shall later urge).

I shall deal shortly with the question: why believe that the Sat predicate is well determined? But first, note that it is this question that should occupy a philosopher concerned with a "scientific" definition of truth, and not the question: Can we eliminate semantical notions? What Tarski showed in CTFL was that semantical notions, for certain languages, could be done away with in favor of set theoretic predicates along with the primitive predicates of the language in question.

Our confidence that these set theoretic and primitive predicates are well determined, transfers onto the notions of truth and satisfaction, which are explicitly defined in terms of these predicates. But there should be nothing suspect about semantical notions that cannot be eliminated, so long as we have some sort of guarantee that the notions are well determined: defining semantical in terms of set theoretic and certain primitive notions is merely a special way of obtaining this guarantee.

We can show that the predicate Sat is well defined, if we permit mathematical induction on the Sat predicate. To allow this, however, is somewhat problematic, for induction on a predicate is usually warranted only on predicates already secured to be well defined. Now I think there is a circle here, but a benign circle. It will be helpful here to present in some detail this proof that the predicate is well defined.

The Sat predicate is intended to be implicitly defined by the axioms in which it occurs. Ordinarily, to demonstrate that a predicate occurring in such axioms is indeed implicitly defined, one presents a certain kind of uniqueness proof. First, one shows that the predicate has at least one extension, and then that it has at most one extension. But showing a predicate has an extension usually comes down to showing there is a certain collection. This collection is such that, if its members are taken to be precisely those objects that

satisfy the predicate, then the axioms in which the predicate occurs are true. Likewise, proving that a predicate has at most one extension is to prove there is at most one such collection. In the case of the Sat predicate presently under examination, however, neither of these things can be demanded: for precisely what is at dispute is the existence of these collections. This does not mean there is nothing left to hope for. One can still prove that if Sat_1 and Sat_2 are two predicates that satisfy the axioms set forth on page 20, then (x) (y) ($Sat_1(x,y) \leftrightarrow Sat_2(x,y)$). One can prove also that whatever we would want to satisfy the Sat predicate does. Formally, this latter is just to show the Tarski biconditionals for all formulas of ZF are provable. Let us turn to these proofs.

First, to show that (x) (y) ($Sat_1(x,y) \leftrightarrow Sat_2(x,y)$).

We use induction on the complexity of the formulas ϕ . Suppose ϕ is of the form ' $x_i \in x_j$ ' and $sat_1 ('x_i \in x_j', s)$; since axiom (1) on page 20 holds of $Sat_1, (s)_i \in (s)_j$. But this same axiom holds of Sat_2 also; hence $Sat_2 ('x_i \in x_j', s)$. Symmetrically, if $Sat_2 ('x_i \in x_j', s)$, then $Sat_1 ('x_i \in x_j', s)$. The inductive cases run just as smoothly. If $\phi = \neg \psi$, $Sat_1 ('\neg \psi', s) \leftrightarrow \neg Sat_1 ('\psi', s) \leftrightarrow \neg Sat_2 ('\psi', s) \leftrightarrow Sat_2 ('\neg \psi', s)$
 If $\phi = \psi_1 \vee \psi_2$, $Sat_1 ('\phi', s) \leftrightarrow Sat_1 ('\psi_1', s) \vee Sat_1 ('\psi_2', s) \leftrightarrow Sat_2 ('\psi_1', s) \vee Sat_2 ('\psi_2', s) \leftrightarrow Sat_2 ('\phi', s)$
 If $\phi = \exists x_i \psi$, $Sat_1 ('\phi', s) \leftrightarrow \exists s' \frac{s}{2} Sat_1 ('\psi', s') \leftrightarrow \exists s' \frac{s}{2} Sat_2 ('\psi', s') \leftrightarrow Sat_2 ('\phi', s)$

The demonstration that the Tarski biconditionals for

all formulas are provable falls out directly from the inductive axioms for Sat. For atomic formulas, the Tarski biconditionals are provable, because it is precisely this that (1) states. The inductive cases are only slightly more involved. Let us handle just the quantifier case. So

let $\phi = \exists x_i \psi$. Then, by inductive hypothesis, $\vdash \forall s (\text{Sat}(\ulcorner \psi \urcorner, s) \leftrightarrow \bigwedge_{\lambda} \psi(s_{j_1}, s_{j_2}, \dots, s_{j_n}))$

So, $\vdash \forall s (\exists s' \stackrel{\sim}{=} s \text{Sat}(\ulcorner \psi \urcorner, s) \leftrightarrow \exists s' \stackrel{\sim}{=} s \psi(s_{j_1}, s_{j_2}, \dots, s_{j_n}))$. But this

is equivalent to: $\vdash \forall s (\exists s' \stackrel{\sim}{=} s \text{Sat}(\ulcorner \psi \urcorner, s) \leftrightarrow \exists x_i \psi(s_{j_1}, s_{j_2}, \dots, x_i, \dots, s_{j_n}))$

where $i = j_\ell$, some $\ell \leq n$. But by axiom (4), $\vdash \text{Sat}(\ulcorner \exists x_i \psi \urcorner, s) \leftrightarrow \exists x_i \text{Sat}(\ulcorner \psi \urcorner, s)$

Hence $\vdash \forall s (\text{Sat}(\ulcorner \exists x_i \psi \urcorner, s) \leftrightarrow \exists x_i \psi(s_{j_1}, s_{j_2}, \dots, x_i, \dots, s_{j_n}))$, which was to be proved. By induction, we may now conclude that all Tarski biconditionals are provable.

The inductive proofs above are quite trivial, just as one would anticipate, for they reflect exactly the inductive definition of Sat. As I have remarked, ordinarily one would hesitate to admit induction on a predicate not known beforehand to be well defined; but in the case of a predicate inductively defined, it does sometimes seem permissible. For accepting the inductive definition of a predicate and accepting these particular inductive proofs using that predicate seem to be two aspects of the same intuitive insight. This insight is the recognition that, if a predicate is inductively defined, then there is just one predicate thus defined.

At all events, the proofs above do at least serve as some kind of formal reassurance that the Sat predicate is well

defined. Certainly, things would be less propitious if no such proof were forthcoming. In the case of ' $R(x,y)$ ', there is indeed no such proof to be found, and our confidence that the predicate is intended "uniquely" by us is thereby undercut, along with any belief that induction involving this predicate is appropriate.

What does Parsons have to say about the issue: how can we justify mathematical induction involving the Sat predicate?

What seems to be required here is that definitions by induction on the natural numbers should be understood and accepted directly or explained by an argument not of a second order character. The first course is certainly conceivable and seems a reasonable course in dealing with a single inductive definition such as that of satisfaction. However, it renounces the attempt to state the principles involved, and it is hard to see how to do that without quantifying over properties or classes or related entities such as propositions or proofs. It seems to me that there might be some alternative that would use the notions of meaningfulness and truth in a way different from the usual uses in formal semantics, in that their extensions would be gradually constructed rather than being definite for a given context.⁷

This passage is not without its difficulties in interpretation. The penultimate sentence in particular is ambiguous. When Parsons writes 'it is hard to see how to do that . . . ' does 'that' refer to the renouncing of the attempt to state the principles involved, or the attempt to state the principles involved? I suspect that Parsons intends to say only that the attempt to state the principles involved requires talk about properties, classes, propositions, or proofs. But then Parsons must be taking it as obvious that to renounce the attempt to state the principles involved is to do something

inappropriate; for at no point does he take further account of the possibility of justifying inductions directly. Since an obligation to state the principles involved seems so little obvious to me (indeed I think that, strictly speaking, there are no such principles), I believe it best at least to consider the alternative interpretation of the sentence. So, for the moment, let us take the sentence in this alternative way.

On this interpretation, Parsons thinks that the obstacle to accepting directly the inductive definition of satisfaction (or, presumably, the inductive definition of anything) is the difficulty expressing certain principles which are implicitly renounced. It is not clear, however, that in accepting directly induction on natural numbers, we are somehow constrained by that very act to renounce the sorts of principles Parsons claims we are. Presumably, in directly accepting induction on natural numbers for particular predicates, we are doing something entirely positive. We are not in the act itself claiming implicitly or otherwise that, e.g., we cannot state that induction is allowed on any class (in the sense of NB or NB+), as Parsons seems to intimate. Indeed there is nothing about our position to require us to make such a claim at any point. Quite to the contrary, our position would be one in which we would not talk about classes (or "related entities") at all. Rather we would largely restrict our comments to those things which we allow to exist, viz.,

sets. If we were ever to talk about classes, it would be in the same spirit that one might talk about Pegasus or phlogiston even when one does not believe such things exist.

It is especially strange that Parsons should think that we must renounce the attempt to state certain general principles in accepting induction directly because he himself shows very effectively that there is no formalization that fully captures our intuitive idea of mathematical induction. If there is no formalization, of any order, which can capture all of the intuitive idea, then there is no general principle, one would think, that would express mathematical induction as we intuitively understand it. But then in recognizing that this is so, we do not somehow embrace or even fail to embrace a general principle which supposedly would capture mathematical induction; we are in the course of perceiving that there is no such general principle. We cannot be renouncing the attempt to state the principles involved when our very stand is that there are no general principles involved, hence none to renounce the attempt to state. Our intuitive concept of mathematical induction is openended; given any formulation, there is a property not expressible in that formalism for which we would also allow mathematical induction. Granted this, we must accept the different versions of mathematical induction theory by theory, not all at once. When we license mathematical induction on the Sat predicate directly, we are merely taking

one of the never ending steps in this theory by theory enrichment of induction.

In the final sentence in the passage from Parsons, he suggests there may be a way to justify induction without second order reasoning: evidently, by somehow having the extension of the truth predicate gradually constructed; Parsons says nothing more about what this possibility might come to. I confess to being unable to understand this alternative well enough to pursue it further. For my purposes however, it is not important that I do pursue it. It is enough that Parsons' misgivings about accepting induction directly do not appear well founded. For the "justification" I have set forth for induction on the Sat predicate is just that it is seen to be warranted directly, in a sense. As I have already observed, however, this direct warrant is unusual because it requires the Sat predicate to be well defined, and, for a formal demonstration of this latter fact, induction on the Sat predicate must be employed. It is indeed the very presence of a circle here that inclines me to say that we accept the legitimacy of both the inductive definition of Sat, and of the use of induction on Sat directly.

We have not so far been confronted by a knockdown argument that shows induction on a predicate cannot be endorsed directly. But maybe to expect this is somewhat to miss the point. Perhaps the idea is rather that there is a certain artificiality in directly accepting such induction. Thus,

conceivably, whenever we motivate, in the privacy of our own hearts, first order induction, we invoke the existence of subcollections of the domain of the quantifiers. We may for the outside world expunge all talk of these subcollections; however, to ourselves, in understanding induction, we may mutter: but still they must exist. Now I think it is true that we sometimes do appeal to second order reasoning in order to motivate first order induction; for example first order induction in Peano arithmetic might on occasion be secured in our minds by the recognition that the principle of second order induction in second order PA (PA^2) is legitimate. However, the second order quantifiers of PA^2 range over entities of whose existence we feel totally assured; hence, the naturalness of the transition to second order reasoning does not seem to count for much.

But the motivation of first order induction for the formulas of ZF does not appear to involve appeals to higher order reasoning. To see this, it is illuminating to consider how mathematical induction is proved for the formulas of ZF. The proof that induction holds for all formulas of set theory is brief enough to include here. What is to be proved is that, for any formula ϕ in the language of ZF, $(\phi(0) \wedge \forall x (x \text{ is an integer} \rightarrow (\phi(x) \rightarrow \phi(x+1)))) \rightarrow \forall x (x \text{ is an integer} \rightarrow \phi(x))$. Well, suppose not, for formula ϕ . Then consider $A = \{x \mid x \text{ is an integer} \wedge \neg \phi(x)\}$. By foundation, there must be a minimal x in the sense of $<$ in this set. Since the $<$ relation among

numbers is just the \in relation, we have picked a minimal such x in the sense of $<$ also. Clearly this x is not 0, since by hypothesis $\phi(0)$. Hence, $x = y+1$ for some y . Since $y < x$, y is not in A . So $\phi(y)$. But $\phi(y) \rightarrow \phi(y+1)$. So $\phi(x)$. Contradiction, since $x \notin A$.

Is there something wanting in this proof? If, in some corner of our minds, we always sought second order reasoning when induction was being justified, we would surely not be altogether comfortable with the proof just as it stands. For the proof has an entirely first order nature. One possible point of weakness in the proceedings might be simply this. The first order principle of induction, even to be asserted, requires the assertion of an infinite number of formulas at once, one induction matrix for each formula of set theory. Being finite beings, how can we do this? And if we can do it, is it not because we see each of these instances as following from the second order principles of induction, which, of course, can be stated in a single sentence? I fail to be convinced by this line of argument. (G. Boolos has an unpublished paper that addresses, among other things, an issue related to this sort of issue. My objections on this score owe much to his arguments there.) Consider the first order schema $(\forall x)(F_x \rightarrow F_{x+1})$. Suppose I have a particular language at hand, and I wish to assert at once all of the instances of that schema in that language. I know what all the formulas are: I know what it

would be and how to assert each individual instance; I also desire to assert each instance, and, finally, to assert all the instances at once. What could prevent me from doing this last?

I may further recognize that $(\forall x)(Fx \rightarrow Fx)$ is valid. Does my insight that this is so depend in any way on an implicit acknowledgement that $(\exists F)(\forall x)(Fx \rightarrow Fx)$ is valid? I see no reason to think it does. For there is nothing peculiarly collection-theoretic in my intuition that $(\forall x)(Fx \supset Fx)$ is valid. Thus, it is not that I first set before my mind all subcollections of the diverse domains for the variable, and then observe that if x is some such collection, it is in that very same collection. Now if there were some use of the combinatorial nature of such collections tacit in my intuition that $(\forall x)(Fx \rightarrow Fx)$ were valid, then I would indeed be resorting to the second order version to underpin this intuition. However, no such use slips in, as far as I can see. Rather, there is nothing more involved than the characteristically boolean properties that predicates have; in this case, the property is particularly trivial: viz., if a predicate holds of an object then it holds of that object. It is true that recognizing this schema to be valid may oblige us to make an extensive survey, of an infinite number of first order predicates, and perhaps this is thought to be problematic. But it is no more problematic than the alternative, which is

to survey an infinite number of subcollections of the domains.

Let us return to the case of induction for the language of ZF, and note the parallels. We of course have a clear idea of what a formula of ZF is; we understand and wish to assert the induction matrix for each formula, for we see they each hold. Why can't we assert them all simultaneously? There should be no difficulty here, if there was none for $'(x)(F_x \rightarrow F_x)'$. What, then, about motivating first order induction: does that involve the second order version? As with the intuition that $'(x)(F_x \rightarrow F_x)'$ is valid, there appears to be no essential use of collection theoretic reasoning. For each individual formula of set theory, we can run through the proof stated above, and see that it applies, and that therefore the induction matrix for that formula holds. But we can also survey all these proofs at once; or at least we can if we can survey all of the relevant subcollections of the domain, as we would have to if we were to appeal to second order induction.

'All this may be well and good,' one might demur, 'to motivate the use of induction for formulas of ZF. But these formulas are significant independently of the use of induction on them. The Sat predicate is not so blessed; precisely its significance is in dispute.' Now certainly there is this important difference between motivating induction formulas of ZF and doing it for the Sat predicate. But the point is

that there is no precedent for turning in the direction of second order reasoning to motivate induction on the Sat predicate; hence we should not apriori expect the justification to come from that quarter.

Of course, the Sat predicate is special, being inductively defined in the way it is, and this may call for second order reasoning in its case, but the intuition underlying our endorsement of the Sat predicate as well defined does not appear genuinely collection-theoretic. Here again, no strictly combinatorial principles seem to intrude. To secure this claim, consider a typical inductive definition.

Suppose I inductively define formulas in the propositional calculus. I say that 'p', 'p₁', 'p₂',... etc. are all formulas; further, if ψ_1 , and ψ_2 are formulas, so are ' $\neg\psi_1$ ', ' $\psi_1 \vee \psi_2$ ' and ' $\psi_1 \& \psi_2$ '; finally, nothing is formula unless obtained by one of these steps. To understand this definition, we may imagine what may be metaphorically described as a certain infinite process. This process begins with the sentence letters 'p₀', 'p₁', ..., etc.' from these, it goes on to create new formulas, e.g., ' $(p_{12} \vee p_5)$ ', ' $\neg p_{47}$ '; from this stage, it advances to generate still more formulas, e.g., ' $(\neg p_{47} \vee (p_{12} \vee p_5))$ '; and so forth. Now suppose we come across some object, say ' $(\neg p_{16} \vee (p_0 \vee))$ ', and we want to determine whether it is a formula. First, we check that it is built up from a finite number of occurrences

of ' p_0 ', ' p_1 ', ..., ' v ', '&', '-'. It is; so we check whether it is ever reached in this infinite process. Well, we may reason, if it were reached at some stage, then, since ' v ' is the main connective, at the previous stage ' $\neg p_{16}$ ' and ' $(p_0 v)$ ' must have been established to be formulas. But let us examine ' $(p_0 v)$ '. At what stage could this have been reached? As before, formulas with ' v ' as the main connective must have a wff on either side of them ' $(p_0 v)$ ' does not. Hence this expression is not a formula; but then neither is ' $(\neg p_{16} v (p_0 v))$ '. We observe that this method applies for any object, so that it will determine whether or not that object is formula.

Now at what point, if any, is collection-theoretic reasoning employed in this definition, or in seeing that the definition works? Let us examine the important steps. No assumption is made that, at the start, ' p_0 ', ' p_1 ', ' p_2 ', etc., are all contained in some collection. The definition could just as easily have begun by letting, say, each of the ordinals be a sentence letter; this sequence we would have comprehended quite as well. When we set before our mind this infinite process, it is not that we must suppose there is a collection that corresponds to the completion of this infinite process, gathering up all the formulas generated along the line. It seems unnecessary also that at each stage on the way there be a collection that contains all the formulas

generated so far. Indeed, it strikes me as no more obligatory to suppose that for each of these stages, or for the completion of all the stages, there be a collection of the formulas obtained at those points, than it is to assume that the iterative process in set theory must be captured by some collection.

Finally, consider the insight that the definition works. Take an arbitrary object. Whether or not it is built up of a finite number of occurrences of ' \neg ', ' \forall ', ' $\&$ ', ' P_0 ', ' P_1 ', ' P_2 ' etc., is surely not something that entails second order reasoning. If it is so constructed, the steps required to reduce it either to subparts that are not formulas, if it is not a formula, or to sentence letters, if it is, invoke nothing collection-theoretic. And neither, it seems, does the insight that this method does the job for any object.

Now the inductive definition of Satisfaction is really no different from this in its motivation. The first step differs somewhat, inasmuch as there are countably many sentence letters, but more than set many pairs for which the Sat predicate holds. But this inessential divergence can be remedied by taking, as I suggested, all ordinals as sentence letters in the inductive definition of a formula.

Enough, then, of all this. Perhaps there are other reasons for thinking that accepting ZFT is tantamount to accepting NB or even NB+. I shall present another possible defense of the view, but it will take some work to motivate. We know that for any first order theory which is consistent, there is a

model in the natural numbers; this is in essence the content of the Lowenheim-Skolem theorem. But if this is so, how is it that we succeed in picking out sets as the domain over which our quantifiers range, and not numbers as this model would provide? Moreover, why is it that we do not take the Lowenheim-Skolem theorem as a reduction of sets to natural numbers? An obvious response is Quine's: we do not effect a genuine reduction because there is a serious price paid in ideology for the savings in ontology. No matter which model of ZF with its domain consisting of the natural numbers we choose, we have the following predicament. There will be relations among the numbers that are, in these models, the interpretations of formulas in the language of ZF, yet are not expressible by any formula in the language of arithmetic. Thus, granted that the ideology of ZF must be considerably more powerful than that of arithmetic, why think that we can get away with a weaker ontology merely by pointing to the conclusion of the Lowenheim-Skolem theorem? Isn't to embrace the stronger ideology, in effect, to embrace the stronger ontology? Quine puts the point this way:

Blanket pythagoreanism on these terms is unattractive, for it merely offers new and obscurer accounts of old moves and old problems. On this score again, then, the relativistic proposition seems reasonable; that there is no absolute sense in saying that all the objects of a theory are numbers, or that they are sets, or bodies, or something else; this makes no sense unless relative to some background theory. The relevant predicates--"number", "set", "body", or whatever--would be distinguished from one another in the

background theory by the roles they play in the laws of that theory.⁸

Quine himself insists that a proxy function be available for there to be a genuine ontological reduction, and this clearly is not present in the case of the purported reduction of sets to numbers. (There are reasons to doubt that this requirement is sufficient, however, as I shall argue later in the paper.)

The same sorts of considerations that vitiate any attempt to reduce sets to numbers might be adroitly parlayed to try to show we are committed to NB or even NB+. For the ideology of ZFT, to which we are committed, clearly outstrips that of ZF by itself. There are relations expressible in ZFT not expressible in ZF alone; viz., those which involve the notions of truth or satisfaction. Granted this increase in ideology, it is not implausible that it actually commits us to a stronger ontology. And since there is a translation from NB(NB+) into ZFT, embracing the ideology of ZFT is in effect to embrace that of NB(NB+). But if we have NB(NB+) as our background theory, or something that it is translatable into, namely ZFT, we appear to be embroiled in the universe of NB(NB+). For, to employ Quine's point, there appears to be no absolute sense in saying that there are, or are not, proper classes; from the standpoint of NB(NB+) as a background theory, there of course will be.

The reader has surely already noted the obfuscation present

in this argument as it stands. In essence, the argument proceeds thus: to be committed to the ontology of a theory is nothing more than to be committed to its ideology; we are committed at least to the ideology of ZFT; but NB(NB+) is translatable into ZFT: therefore, we are committed to those entities to which NB(NB+) would commit us. However, the argument elides the critical question: given that NB(NB+) and ZFT are translatable, the first into the second, in the sense in which they are, which ontology should we accept? Evidently, the idea behind the argument is that, if we accept a theory, we are committed to all the entities that that theory, or any theory translatable into it, alleges that there are.

At first blush, this might appear to fly in the face of what goes on in the case of ontological reduction. For presumably when we reduce one ontology to another there are two theories, one of these translatable via a proxy function into the other, and the former theory claims there are certain things the latter does not. We consider this translatability of the one into the other to show that we can dispense with the one ontology in favor of the other. But in fact, it is not clear that this is the best way to understand ontological reduction. Let us consider a paradigm case of ontological reduction, the reduction of number theory to set theory. Now it will be convenient, for later purposes, to transform the reduction into one between two equivalent theories. So construct

a two sorted theory, in which one kind of quantifier ranges over numbers, and the other kind of quantifier ranges over sets. That part of the theory which has quantifiers over sets would have the power of ZF; the part which has quantifiers over numbers would be of the strength of PA. In the two sorted theory, ' $3 \in 5$ ' for example would not be meaningful; while it would be in the proposed theory to which the two sorted theory would be reduced, ZF. For all that, however, under the usual translation the two theories would be equivalent. So here we might seem to have a case in which there are two intertranslatable theories, one of which holds there are things which the other does not. And we surely consider this an ontological reduction.

But is there good reason to say that the two sorted theory claims that there are certain entities which ZF does not? May we not take the result to be precisely that numbers, over which one set of quantifiers in the two sorted theory range, just are sets? To put it differently, the results might be understood to show that any belief we might have had, while accepting the two sorted theory, that numbers were distinct from sets, is mistaken; in fact numbers are nothing but sets. (Of course, if we had had this belief, it is not something that the two sorted theory alone commits us to or should persuade us of.) The proposal here is not that numbers are objects in the sense of Frege.⁹ Rather, the point is that in committing ourselves to ZF alone,

we are not escaping commitment to numbers. In holding ZF, we are as much committed to numbers as we would be in adopting ZF+PA; this is what the ontological reduction is taken to show.

Assume, then, that ontological reduction between theories involves precisely a demonstration that certain entities of sort A are really certain entities of sort B. Where does this leave us with regard to NB(NB+) and ZFT? It might appear to show that we are not avoiding the ontology of NB(NB+) by accepting ZFT; for the ontology of NB(NB+) might somehow be the ontology of ZFT.

Although the suggestion that NB or NB+ can be reduced to ZFT may seem rather bizarre, it will be instructive to deliberate this question. Aside from its direct consequences for the question of our obligation to admit proper classes, we shall see that, in considering it, there is much to be learned that will bear on what we have already covered. Moreover, while the proposal may appear at best something of a curiosity, it is difficult to abjure on principled grounds. At any rate, this proposal shall occupy us in the last chapter of this thesis.

CHAPTER II
TRUTH THEORIES AND THE HIERARCHY OF V

In "Sets and Classes,"¹ Parsons observes that there is a translation between NB+ and ZFT.* What Parsons calls NB+ is not what is usually intended by the term. Ordinarily, by 'NB+' is meant Kelley Morse set theory: the extension of NB obtained by allowing bound class variables in the class existence axioms. For Parsons, NB+ is the theory got by extending the replacement axioms (for sets) to include those with bound class variables. This is a much weaker theory than Kelley Morse set theory; how much weaker may best be seen by a proof that the translation from NB+ (as Parsons intends the term) to ZFT fails as a translation from Kelley Morse set theory into ZFT. The translation Parsons has in mind is this. Take a formula of NB+. We can consider ' $(\forall Y)$ ' to be defined as ' $-(\exists Y)-$ '. Now wherever $\exists Y(\dots Y \dots)$ occurs, replace it with $\exists_n \exists_s(\dots \{u | \text{Sat}(n, s^o)\} \dots)$ s^o, u is the sequence just like s save that u is substituted at the 0th place. Eliminate the abstract, which is virtual. This formula will be true in ZFT if the original

*By ZFT I mean what I meant in the first half of my paper. Take the language of ZF; add a two place predicate $\text{Sat}(x, y)$. Adjoin the ZF axioms which define inductively satisfaction for the formulas of ZF. Finally allow formulas built up out of $\text{Sat}(x, y)$ to be used in replacement axioms.

formula was true in NB+ (on a standard notion of truth for NB+).

Suppose the same translation worked for Kelley Morse set theory. I claim the following is provable in Kelley Morse:

$$1) \exists x \forall n \forall s - [(\forall u (u \in x \leftrightarrow \text{SAT}(n, s^0, u)))]$$

Here $\text{SAT}(x, y)$ is a predicate of NB+ (hence of Kelley Morse) that expresses satisfaction for formulas in ZF, obtained in the manner Parsons suggests. The translation of 1) is

$$2) \exists n, \exists s, \forall n \forall s - [(\forall u (\text{Sat}(n, s, u) \leftrightarrow \text{SAT}^*(n, s^0, u)))]$$

where SAT^* is the translation of SAT . Hence, by quantifier logic there is a n_0, s_0, u_0 , such that

$$3) \text{Sat}(n_0, s_0^0, u_0) \leftrightarrow \text{SAT}^*(n_0, s_0^0, u_0)$$

But consider: in Kelley Morse all the Tarski biconditionals for formulas of ZF are provable, and since Kelley Morse is not ω -inconsistent, n_0 must be the godel number of a genuine formula, say φ . Hence,

$$4) \text{SAT}(n_0, s_0^0, u_0) \leftrightarrow \varphi(x_0, x_i, \dots, x_k / u_0, s_i, \dots, s_k)$$

is provable. But its translation 5) should then be provable in ZFT.

$$5) \text{SAT}^*(n_0, s_0^0, u_0) \leftrightarrow \varphi(x_0, x_i, \dots, x_k / u_0, s_i, \dots, s_k)$$

Yet in ZFT, all Tarski biconditionals are provable, including

$$6) \text{Sat}(n_0, s_0^0, u_0) \leftrightarrow \varphi(x_0, x_i, \dots, x_k / u_0, s_i, \dots, s_k)$$

But 5) and 6) give us

$$7) \text{SAT}^*(n_0, s_0^0, u_0) \leftrightarrow \text{Sat}(n_0, s_0^0, u_0)$$

which contradicts 3).

It remains to show that 1) is provable in Kelley Morse. In essence this is just an application of Cantor's proof that there are more members in the power set of a set than in the set itself.

Proof of 1). Suppose to the contrary that

$$8) \forall X \exists n \exists s [(u) (u \in X \leftrightarrow \text{SAT}(n, s^{0,u}))]$$

Define A thus:

$$9) A = \{ \langle n, s \rangle \mid \exists Y (u) (\text{SAT}(n, s^{0,u}) \leftrightarrow u \in Y) \& \langle n, s \rangle \notin Y \}$$

A must exist in Kelly Morse. By our hypothesis, there is a n_1 , and a s_1 such that

$$10) \forall u (u \in A \leftrightarrow \text{SAT}(n_1, s_1^{0,u}))$$

Suppose $\langle n_1, s_1 \rangle \in A$. Then $\exists Y [\forall u (\text{SAT}(n_1, s_1^{0,u}) \leftrightarrow u \in Y) \& \langle n_1, s_1 \rangle \notin Y]$.

But any such Y must, of course, be identical to A; hence

$\langle n_1, s_1 \rangle \notin Y$. Suppose now that $\langle n_1, s_1 \rangle \notin A$. Well then 8) must be true by EG on A. But then $\langle n_1, s_1 \rangle \in A$. Contradiction.

We can see how far short of Kelley Morse set theory is ZFT (and NB+). Now let us consider the following. Suppose we start out with ZFT, recognize its equivalence with NB+, and then go on to adopt NB+; next we establish a truth theory for NB+, recognize its equivalence to a further super-class theory, and suppose we iterate this process as often as we see fit. What will the structure we get by these means look like? We shall see that it bears no great resemblance to a continued iteration of the ranks of V. Reflect on the implications of this fact. Parsons' argument that we cannot quantify over all sets might be put thus. We seem to need a truth

theory for ZF; but this truth theory requires for its justification (or is equivalent to a theory which includes) proper classes. But these proper classes are similar to a continuation of the iterative hierarchy. Yet once we introduce new ranks in the iterative hierarchy, we should continue as the axioms of replacement and power set would require. Isn't the most plausible way to make sense of this situation to hold that our quantifiers never really range over all sets, but only over V_α for some α ?

Parsons' argument is seriously undermined if we see that the justification for a truth theory, even if it involved us in proper classes, and superclasses, etc., at no point involved us in a commitment even to one additional, complete rank of collections. For then there is little temptation to see the level of proper classes as a continuation of the iterative hierarchy. And certainly there will be no impetus to start applying the axiom schema of replacement if even the level of proper classes is not a plausible candidate for having been obtained (in part) by a full blooded separation axiom. Let me explain this last point. The principle which most clearly exhibits the combinatorial feature of collections is the axiom schema of separation (or, in a more powerful way, the axiom schema of replacement). The power set operation by itself does not really provide us with the combinatorial aspect, since the set of all subsets

may be very small if we don't permit strong principles for separating off subsets. Rather, the power set axiom is best understood as merely a principle for iterating new levels of sets. The axiom of separation arises from the idea that a subset of a set should be held to exist, no matter how the "decisions" to include and exclude members of the original set are made, so long as the decisions are made for all members, and the decisions are consistent (that is, there are not two decisions, one to throw the member in, and one to throw it out). The axiom of replacement is motivated by a kind of extension of this reasoning.

It has, I think, been insufficiently recognized to what extent ZF embodies an iterative-combinatorial conception of set, and not just an iterative conception of set. Thus as I have said separation and replacement can be justified on combinatorial grounds but not on purely iterative grounds. In addition the axiom of choice, while of course independent of ZF, purports (rightly, I think) to be justified combinatorially, and again is not justified by the iterative conception alone. Similar points hold, I believe, for higher axioms of infinity, e.g., the existence of a strongly inaccessible cardinal. But more on this in the next chapter of my thesis.

Now the extension of ZF, described above by proper classes, then by super-classes, then by super-super-classes,

etc., will never give us a full rank of collections at the level of proper classes. What I mean by a full rank of collections is that at least the existence of impredicative classes should be provable; that is, that at least Kelley Morse set theory be provable. Given that this extension is so impoverished combinatorially by comparison to Kelley Morse set theory, we have little reason to conclude the quantifiers of ZF should be understood to range over a set. For the next rank above the level of the quantifiers would in this case include even impredicatively defined subsets of the set which is the range of the quantifiers; the combinatorial feature of sets here would require that such subsets would exist. But there is no impetus to think that the collections we get by the iteration above described are so closed, for their motivation is entirely different in character; they need only provide a backdrop against which certain truth theories may be developed. Since the combinatorial feature of sets is an entirely central one in our conception of the sets in ZF, there seems to be no absolutely compelling ground, deriving from our acceptance of this hierarchy, to believe that the range of the quantifiers of ZF should be construed as a set.

Let us see more technically what my claims come to. Precisely, the third order theory we might adopt, in order to have a theory that does for NB+ what NB+ did for ZF, is this. The new, additional class axioms are all instances

of $\exists x^2 \forall y^1 (\gamma \in x^2 \leftrightarrow \varphi(y^1))$ where φ is any formula of NB+, x^2 is a third order variable, and y^1 is a second order variable ranging over the classes of NB+. Now, if we are going to prove that induction on formulas of this new theory works, as we must if we are to prove, e.g., all provable formulas of NB+ are true, then we must add the following replacement axioms.

$$\forall x \exists! v \varphi(x, v) \rightarrow \forall u \exists v \forall x \in u \forall y (\varphi(x, y) \leftrightarrow y \in v)$$

for all formulas φ of the new third order theory. Note that we have this replacement axiom only for sets. We could have thrown in a (kind of) replacement axiom (which just amounts to a separation axiom schema) for proper classes as well, thus:

$$\forall x \exists! v \varphi(x, v) \rightarrow \forall u \exists v \forall x \in u \forall y (\varphi(x, y) \leftrightarrow y \in v) \\ \text{(or, equivalently, } \dots \rightarrow \exists v \forall x \forall y (\varphi(x, y) \leftrightarrow y \in v) \text{)}$$

where φ

is a formula of the new theory. This axiom implies, of course, Kelley Morse set theory. But we do not have to add this axiom to prove the semantical facts we set out to prove in the first place; for the syntax of NB+ can be completely coded up in ZF (in fact, in the set of hereditarily finite sets). For this reason, to show that e.g. all provable formulas of NB+ are true we need only the axiom of replacement for sets (indeed, we need only the related axiom of separation for subsets of ω). If we were to adopt this axiom, it would have to be for reasons other than those which originate from

our desire for semantical theories.

We may develop fourth order, fifth order, etc., theories analogous to this third order theory. A natural question to ask is: how high up do we want to have these theories to be iterated? Presumably, as high as we want to accept the corresponding truth theories. But how high up is that? In the first half of my paper, I was cagey about this issue; I intend to remain cagey. However, let me note some extenuating circumstances for being this way. To begin with, it is not clear that this question admits of a definite answer; nor is it clear that if it did have a definite answer that we would ever be able to know it. Since it does appear that if we accept a given truth theory, we will accept a truth theory for that theory, a definite answer to the question must take the form of a least upper bound on the levels of the truth theories we should (or might) adopt. Now perhaps there is no way we could give the expression 'theories we should (or might) adopt' a content determinate enough that such a bound could reasonably be thought to be fixed. But suppose we did think such a bound were determinate. We are all familiar with the arguments that we may be like certain Turing machines, in the theorems we can prove, but cannot know which Turing machines we are like.² A similar argument might show that, though there may be a definite answer to the question of how high up we

can go with our truth theories, we can never know that answer to be the answer.

One might hope that, even if there is no least upper bound, or if there is one but we cannot know it, we can at least know some upper bound. The level of the first uncountable ordinal might appear to be such a level. For it might seem that we cannot continue beyond a countable level, because the natural way of doing this would require us to embrace a theory with an uncountable language. But we can understand only countable languages; hence the relevant truth theory would be in an important sense unintelligible to us. (The truth theory would have to have an uncountable language if all the Tarski biconditionals for the languages below the uncountable level are to be provable.) I am not entirely convinced, however, that we cannot understand an uncountable language as neatly formulable as this.

But in any case, it is fair to say that we cannot go to a level so high that (speaking somewhat loosely) there are more (cardinally) such levels than there are ordinals in the range of the ZF quantifiers. Not, at least, unless we have already determined there are more (cardinally) ordinals than there are ordinals in the range of the ZF quantifiers. For if we have never had any reason to believe that there are cardinally more objects in the universe than there are ordinals in the range of the quantifiers of ZF,

what reason could we have to iterate the truth theories more times than there are ordinals in ZF?

At this point, observe what Kelley Morse set theory does in one fell swoop: provide an ontology of objects greater in cardinality than all the objects in ZF. It was essentially a proof of this fact that showed the would-be translation of Kelley Morse into FOT failed. It is fairly easy to see that, since we can iterate these theories only as many times as there are ordinals in ZF, these class theories will never commit us to more than the number of objects in the universe of ZF.

These cardinality considerations suggest quite powerfully how different is a commitment to the sorts of class hierarchies that truth theories might seem to require from a commitment to just one more full level in the iterative-combinatorial hierarchy.

Another way to see the disparity between the combinatorial notion of collection and an essentially predicative notion is to consider the following possibility, set forth by Parsons. Suppose we take ZF, extend its language to include new primitive predicates, and permit replacement for formulas built up out of these predicates. We can then set up a class theory which assumes the existence of a class corresponding to the interpretation of each new primitive predicate, and to the interpretation of each new compound predicate got by first

order operations. Parsons says that unless we are engaged in 'obvious cheating', such as postulating for each impredicative class a new primitive predicate that will have that class as its extension, he does not see how we could get thus all impredicative classes.

Now the proof of the failure of the translation between Kelley Morse and ZFT can be extended to a proof of something much stronger; it can be utilized to show how obvious such cheating must be: even if we allow as many new primitive predicates as there are objects on ZF, and each of them has a distinct interpretation, we cannot get from them all impredicative classes. Hence, to cover all impredicative classes in the manner Parsons outlines, one would have to assume the existence of a syntax as grand in proportions as the level of impredicative classes would be. What is specially noteworthy about this fact is how quickly the attempt to develop, predicatively, an impredicative class theory runs aground, for one might imagine that it would be in the realm of semantics that problems would first arise, in particular with regard to the question of how we could succeed in intending the primitive predicates so that their extensions taken together, covered all impredicative classes.

Despite all I have shown, one might still argue thus. The best way to take the new levels of collections in the

theories described before is to see them to be continued levels in the iterative-combinatorial hierarchy. Merely because we cannot prove that any of these levels are full levels does not mean we cannot extend our theories and take them to be such. And taking the range of the quantifiers of ZF to be a set seems neater than to adopt this somewhat repugnant hierarchy of stunted ranks.

There is some merit to this argument, I think; but I shall use the argument to opposite effect. For I intend to show that it is not reasonable to commit ourselves to proper classes of any kind, not even those of NB. But this I will do in the last chapter of my thesis.

CHAPTER III

For years, W. V. Quine has held that in view of Russell's paradox, no one set theory should enjoy pre-eminence. It is something of an embarrassment to Quine's position that logicians are, on the whole, seriously interested in just one set theory, ZF (and its conservative extension, NB). One naturally expects the logicians' fascination to be explained by a particularly compelling notion of set embodied in ZF. And one is not disappointed. This notion of set has been described as the iterative conception of set. Now while I agree that ZF expresses an iterative conception, I am persuaded there is another aspect of the sets of ZF not aptly depicted as iterative. Following Bernays,¹ I shall call this aspect the combinatorial. On iterative grounds alone, the replacement axiom, the separation axiom, the axiom of choice, and possibly even the axiom of union are problematic. But the combinatorial feature straightforwardly justifies these axioms.

Two axioms, the axiom of foundation and the axiom of power set express that aspect of sets that I call iterative. The power set axiom by itself epitomizes a very important proper part of this conception, what one might term quasi-iterative. The axiom of foundation constrains how it is that sets can come to be. The core idea may be put in a metaphor: if we can discern a set, we can set up a ladder to get to it.

That is, metaphor excised, all the members of a set are ontologically anterior to the set itself. Hence there can be no set that has itself as a member, no set one of whose members has it as a member, no set one of whose member's members has it as a member; etc. More strongly, there is no set with an infinitely descending \in -chain, so that $\dots \in X_2 \in X_1 \in X_0$. This last principle is reasonable, since if a set depends for its existence on its members, it seems unacceptable to shift the burden without end: in time we must come to an object that stands on its own. The axiom of power set does not impose constraints, but rather provides for the creation of new sets from old; if we have a set, we can get from it the set of all its subsets. Significantly, of these two principles, the axiom of foundation was the later to be conceived and admitted in the development of ZF. It is significant because the axiom is purported to be the most decisive in ruling out Russell's paradox, and is almost the soul of the iterative conception. One can only wonder: if this axiom is so much what ZF is about, what picture stood behind the set theory proposed by Zermelo, before this axiom was even thought of? For Zermelo's theory was only quasi-iterative: it sanctioned the generation of powerful new sets from given sets, but offered no explanation of how the given sets came to be given. Moreover, in an even earlier prefiguration of ZF, the partial system Cantor set out in his 1899 letter to Dedekind, not even

the power set axiom is included. Such considerations set one directly to muse about what is really afoot in ZF. In this doubtful state, it behooves us to look with care at the historical development of ZF. For this history well reflects the intuitive motivation of the conception of set in ZF.

One learns in philosophy classes, and in mathematical logic classes as well, that it was Russell's paradox that laid to rest the "logical" concept of set. On this concept, for any predicate, there is a set that is its extension. For Frege's attempt to reduce mathematics to logic, Russell's paradox was catastrophic; for though perhaps arithmetic did not really totter, Frege's program certainly did. This confrontation between Frege's would-be reduction and Russell's elegant paradox is fascinating--indeed too fascinating. For the thrall it has exerted over philosophers of mathematics has tended to obscure the paradox that I believe actually lies behind the development of ZF. I am talking about the Cantor paradox.

It should come as no surprise that it should be Cantor who first saw what direction set theory should take, and that he should isolate the paradox that was the source of the difficulty with the "logical" notion of set. For fundamentally Cantor was not engaged in some program in epistemology and metaphysics, as was Frege; rather, he was trying to make out what sets were. While the logical notion of set is rightly

so called in the case of Frege, since he was preoccupied by logical principles, for Cantor it is more befitting to regard the unlimited comprehension axiom as formulating a naive notion of set. Wrestling with this concept, he saw, quickly enough in view of the subtlety of the issues involved, it was a notion that could not be sustained, and must be replaced by a more sophisticated notion. In reaction to Russell's paradox, there is a tendency, quite natural if one is approaching the unlimited comprehension axiom from Frege's point of view, to ask the following question: Since Russell's paradox has laid low the full comprehension axiom, how must we restrict this axiom to get one that will not burgeon into a contradiction, and yet will be as close an approximation to this axiom as possible? This tendency is natural from Frege's point of view, since for Frege the justification of the comprehension axiom is chiefly logical, and one wants to save as much of logic as one can. Quine, evidently, has inclined to view the situation with the logical (alias naive) notion of set in just this light, and has asked precisely the above question. No wonder, then, that Quine's way out of Russell's paradox in NF is to impose a purely syntactical constraint that seems to let as many predicates appear in the comprehension axiom as consistency will suffer (and, in its first formulation, Quine's constraint allowed more).

For Cantor, the chief antinomy to come to terms with was

not really Russell's, but one just a few steps removed from it: If every set has less cardinality than its power set, then the set of all sets, which is its own power set, must have greater cardinality than itself. A natural response to this antinomy (though probably not to Russell's) is that there cannot be a set of all sets, because there are too many sets. For a condition on a multitude being collected together into a set is that it have a determinate number of members, a fixed cardinality; but precisely this the set of all sets could not have.

Now from Cantor's 1899 letter,² it is clear that the consideration of cardinality seemed to him decisive in determining whether a set exists. In the face of his paradox, Cantor held firm on the following principles: a multitude is a set if it has the same cardinality as a set; all sub-collections of a set exist as sets; the union set of a set exists. These principles show up in ZF, of course, as the axioms of replacement, separation, and union. What, if anything, do these principles share? They are combinatorial.

What is it for a principle to be combinatorial? On the finite level, combinatorics studies the number of elements in certain sets, the relative sizes of various sets, the number of all the possible permutations of a set meeting certain conditions. In a word, combinatorics is concerned with issues of cardinality, particularly as they bear on permutations of a set. The last business with permutations epitomizes one

feature that is to be transferred from the finite case to the infinite. Bernays puts the point thus:

Passing to the infinite case, we imagine functions engendered by an infinity of independent determinations which assign to each integer an integer, and we reason about the totality of these functions. In the same way, one views a set of integers as the result of infinitely many acts of deciding for each number whether it should be included or excluded.³

The axiom of separation is justified on these grounds: given a set, for a particular member x , either $\phi(x)$ holds or it does not; this formula will "decide" for it whether it is to be included. But so also is the axiom of choice. Consider this formulation of the axiom of choice: For any x , if x is a relation with domain u , then there is a subrelation of x , with domain u , which is a function. Why is this at all plausible? We may imagine a certain series of independent "decisions", one decision for each ordered pair in the relation x . In this series, as it turns out, for each first coordinate in some ordered pair in x , there is exactly one ordered pair with that first coordinate decided to be in the subset of x . If decisions to include and exclude members are independent of each other and basically arbitrary, how could it be that there would not be such a series? That we cannot define this series by a separation axiom should not erode our confidence in the existence of such a series.

Now it may seem that to assume such a series is at any rate exactly equivalent to the belief in the axiom of choice,

and so cannot justify such a belief. But this is to misunderstand how this consideration is intended to underpin the axiom of choice. Mark here the analogy with the motivation of mathematical induction. One often hears induction formulated in the following way. If 0 has a property P , and if whenever n has P , $n+1$ has it, then all numbers have P . But in our more reflective moments, we likely wish to rid ourselves of property talk; and we certainly wish to extirpate the assumption that corresponding to every predicate there is a property, for that way lies inconsistency. Yet once we try to cast this principle in some formalized language, we see we can never fully capture the intuitive content of the original principle. Certainly no first order formulation covers all of it, since for any first order language, it is obvious enough that there will be predicates not expressible in that language. And the second order formulation is no better in this respect, inasmuch as there are predicates, in particular the truth predicate of the second order language, which cannot be expressed in the second order language. And so on for even higher order languages. Does this mean that the principle as originally stated cannot serve as a heuristic, motivating in a sense the sundry formal extensions of mathematical induction? We might try perhaps to salvage the full generality of the original intuitive principle, without the imperfections of that principle, thus: For any predicate ϕ in any language we find meaningful,

if $\phi(0)$ holds, and if whenever $\phi(n)$ holds, so does $\phi(n+1)$, then for all n , $\phi(n)$ holds. But this principle is only as transparent as the expression 'in any language we find meaningful', and how to construe talk about all such languages is a problem notorious for its intractability: the crew of heterological paradoxes lurks here. So if we are to find some way to indicate the generality of the principle of mathematical induction, we must settle for something that is heuristic, something schematic.

Now it is just such a role that the talk of independent decisions plays, in suggesting one aspect of combinatorial closure. If such discourse seems more nebulous than the intuitive statement of induction, it is because combinatorial closure is by its very nature a less tidy notion. An intuitive statement of a general separation axiom, e.g., for any property P and any set z , there is a subset of z of precisely those members of z that have property P , is in closer analogy to the case of mathematical induction. But the essentially non-predicative nature of combinatorial closure will not yield a heuristic principle that can be as neatly formulated as that for mathematical induction. Nor does this intuitive statement of the separation axiom indicate its fundamental justification. Such justification derives from the more basic picture that lies behind the axiom of choice also.

Go back, now, to the case of the axiom of choice and the picture of the series of independent decisions that purports to justify it. Asserting the existence of the series of decisions, which was described before and which would justify the axiom of choice is very much like assuming the axiom of choice itself. But the overall picture does give us reason to believe in the existence of the series. For it seems such pictures serve not only to instruct, but also to justify, in that the pictures enable us to motivate the axioms.

I should say a word about the axiom of union. This axiom is one whose ground appears to be as much iterative as it is combinatorial. For if at some rank we have a set, then its members must all have been present at a rank below, and therefore all its members' members present at even lower ranks. If at each new rank we form all subsets of sets in lower ranks, surely the union set will be there! But here, as with separation, the catch is this: how are we to ascertain what all the subsets of a set are? The point of calling this principle is part combinatorial is well exhibited in an independence proof for the axiom of union: for there what the axiom would do is obtain a set of greater cardinality than all the sets of the model. The independence proof is simple enough. Let the model be the set of all sets which are hereditarily less than \aleph_ω in cardinality. Interpreting ' ϵ ' as ϵ , all the axioms of ZF are true in this model, save for the

axiom of union. For the set $\{L_n\}_{n < \omega}$ would be a member of the model, but its union, L_ω , would not.

But what of the axiom of replacement? Here is a sticking point for any purely iterative conception of set. For, as Parsons has remarked, even if we allow that there are ranks as high as each well ordering generated in the iterative conception, we cannot get the generality of the axiom of replacement. In due time, I shall challenge the claim that the iterative conception can by itself support the existence of ranks corresponding to such well orderings. But remark that even on this generous interpretation of the iterative conception, we see we cannot get the axiom of replacement genetically; that is, given that we have iterated the ranks of V up to V_α , there is nothing in the structure of V_α , no encoded "information", that would lead us to conjecture ranks as high as the axiom of replacement would furnish us. The axiom of replacement permits us to focus exclusively on cardinality considerations in postulating further ranks. By 'cardinality' here is not of course meant the existence of a 1-1 function, as an object, but of a functional, that is, a formula $F(x,y)$ of ZF such that $(x) (E!y)F(x,y)$. Otherwise the axiom of replacement would follow trivially from separation.

Since so much of ZF is not accounted for on the iterative conception, should this conception of set be abandoned? No; supplemented with the combinatorial aspect of sets, it

stands. On the combinatorial notion, as I am portraying it, if there are not too many things associated with the members of a given set, then all those things are bound up into a set. And what is a straightforward way of making a restriction on how many things per member there can be? By permitting only one thing associated with each member. The idea behind replacement is in a certain sense the converse idea of the thought the Cantor paradox led to. The Cantor paradox demonstrated there were too many sets for them all to be collected into a set. This principle says that if there are not too many members of a multitude, then that multitude is a set.

Now the combinatorial principles are of course concerned with issues of cardinality, and the axiom of replacement gives guarantees of the existence of certain sets based on considerations of cardinality. This suggests that the axiom of replacement may be classified as combinatorial. But suggestion is not enough; we want more feeling for how this picture works. The reflections below may help here.

The paradoxes occasioned our fall from Cantor's original paradise. After this debacle, Quine has it, we can only strive variously to recapture in our set theories what we can of the formerly exalted status of the naive notion. Now there may be some justice in the feeling that, after the paradoxes, we are seeking to regain what we can, but not able

to get everything. But only one such program appears to be well motivated, that of Cantor and Zermelo. The idea of V may be construed, somewhat quaintly, as a regulative principle; V thus becomes a kind of tower of Babel, if not a paradise itself. Once cardinality was seen to figure so crucially in sethood, the axiom of replacement was a first natural step to restore some of the bygone strength to the new set theory. Other such devices are the sundry higher axioms of infinity that demand increasing height to the universe. These higher axioms of infinity assert the existence of diverse sorts of cardinals, for example, an inaccessible cardinal, a Mahlo cardinal, a measurable cardinal, a compact cardinal. In all such cases, the properties the cardinals are to possess are generalizations of properties that ω has. So far, there are no incompatible large cardinal axioms. But the fact that $V=L$ is inconsistent with the existence of a measurable cardinal intimates this might happen. What this result suggests is that the impossibility of a measurable cardinal in L is due to the narrowness of L , inasmuch as L cannot contain a measure for any cardinal. That is, 'There exists a measurable cardinal' implies a certain kind of breadth to the universe as well as a certain height. There is then the possibility that a new higher axiom of infinity would entail a different, incompatible filling out of the ranks of the universe. Such a possibility indicates that the combinatorial notion of set is not entirely

determinate; but this should not amaze after Godel's incompleteness theorem.

The axiom of constructibility has to most set theorists seemed a very implausible one. But to take such a view is to have a peculiar idea of how the ranks of V must be fleshed out. It is in essence a rejection of the notion that the existence of a set depends in any way upon definability, and is an embracing of the opposing, combinatorial picture, on which sets will exist in as arbitrary ways as we can conceive. Indeed, even the Continuum Hypothesis, which to Cantor and Hilbert appeared clearly true, but just in need of proof, is becoming ever more viewed with scepticism. Such scepticism has its roots in the arbitrariness of how subsets of a set may "come to be". Thus Cohen has this to say:

A point of view which the author feels may eventually come to be accepted is that CH is obviously false. The main reason one accepts the Axiom of infinity is probably that we feel it absurd to think that the process of adding only one set at a time can exhaust the entire universe. Similarly, with the higher axioms of infinity. Now \aleph_1 is the set of countable ordinals and this is merely a special and the simplest way of generating a higher cardinal. The set C is, in contrast, generated by a totally new and more powerful principle, namely the Power Set axiom. It is unreasonable to expect that any description of a larger cardinal which attempts to build up that cardinal from ideas deriving from the Replacement Axiom can ever reach C . Thus C is greater than $\aleph_\alpha, \aleph_\omega, \aleph'_\alpha$ where $\alpha = \aleph_\omega$ etc. This point of view regards C as an incredibly rich set given to us by one bold new axiom, which can never be approached by any piecemeal process of construction.

Perhaps later generations will see the problem more clearly and express themselves more eloquently.⁴

Now Cohen in this passage maintains that the "incredibly rich set C" is given to us by the power set axiom, and of course this is in part accurate. For without the power set axiom the multitude of subsets of a set would not be a set. But the "incredible richness" of this set originates elsewhere, namely in the diversity of ways a subset of a set can come about.

I have spoken here of how sets "come about"; and I have spoken also of series of independent "decisions" which separate off subsets. Such talk may seem to intimate a certain constructive element in our concept of set. But to take such talk this way is precisely to misunderstand it. For a salient feature of the combinatorial notion is just its separation of the conditions of sethood from anything mind dependent: the combinatorial view of sets is the most full-bodied expression of mathematical platonism. That we should revert to anthropomorphic and physical metaphors in our attempt to delineate this concept is not to be marveled at; it is difficult to see how we could at first come to grips with it, without idealizing certain actions and human capacities. Thus the talk of a series of independent decisions is intended to convey how the existence of a set is determined entirely by what obtains "out there". And when I say that a subset of a set comes about by such a

series, I am rather tracing the progress of our recognition that this subset must exist than describing some metaphysical process of generation. There is one point at which it may seem appropriate to depict sets as being generated; this lies in the more strictly iterative aspect of sets of ZF. However, even here it is quite metaphorical to portray it as generative. To this point I will return. In any case, ideally, after getting a handle on the concept of set, our intuitions will be so sharpened that metaphors will be no longer necessary.

The suggestion Cohen makes in the passage quoted links up in a remarkable way two seemingly disparate aspects of the combinatorial notion: the breadth of each rank of V , and the height of V . The proposal would regard the successor ranks as so broad that, speaking loosely, all the power of replacement on ordinals that can be described at that rank could not furnish a cardinal large enough to match the rank itself. Hence, when replacement is used on that rank, as it can be when it is gathered up into a set at the very next rank, it shoots up the height of the universe more than could be dreamt of before.

While Cohen's thought seems prompted by combinatorial considerations, such could also be invoked against it. Thus if there are so very many subsets of a set, won't there be just a tremendous number of functions generated at each rank? And

might not these functions serve to press down the cardinality of the power set of a set? To defend Cohen's view, one must appeal to complicated intuitions about how it is much easier to bring about random sets than functions from sets into very proper subsets. I think such an intuition can be defended. For the combinatorial notion seems to be like a principle of greatest entropy: mathematical objects, that is, sets, are as disorganized and unsystematic in the ways they come to be as possible; and the generation of functions that suppress cardinality seems to imbue this generation too much with a sense of order and method.

A generous reader may forgive this unrefined speculation. At all events,, suffice it here to remark the close tie between the perceived height and the breadth of the universe: for depending upon how we take the ranks to be filled out, the apparent height of the universe may vary dramatically. The interrelationship of the perceived height and breadth is subtle and complicated. Thus, as I have pointed out, the assertion of the existence of a measurable cardinal, while to all appearances an axiom about the height of the universe, would imply something about its breadth. But what is more remarkable is that the axiom requires that 2^{\aleph_1} be a countable ordinal, that is, that there be functions in the universe (obviously not in L) that map ω onto 2^{\aleph_1} . Thus if a measurable cardinal exists, the structure of L is collapsed into a

meagre residuum. So in a sense we have come full circle. An axiom ostensibly about the height of V shows that the ranks of V must have a rich kind of fleshing out. And the sort of breadth these ranks have demands in turn that \aleph_1 in the universe be extraordinarily high, namely, much higher than $|2^\omega|^L$.

Another striking result is this. The axiom of determinacy (AD), an assertion at its base about the power set of ω , that is, about $V_{\omega+1}$, implies properties for cardinals lying far above $V_{\omega+1}$. For if AD, then \aleph_1 and \aleph_2 are measurable, \aleph_n ($n > 2$) are Jonsson, and \aleph_ω is Rowbottom.⁵ Here we see what impact a statement quantifying over the members of a very low rank may have upon the height of V .

Of course, another consequence of AD is that the axiom of choice is false. And it may seem ironic that I should claim the axiom of choice is justifiable combinatorially, when AD is precisely a generalization to the infinite of a principle of finite combinatorics. At this juncture, there is little else to say than that there are more and less credible generalizations from the finite; indeed there are generalizations from the finite that are provably false. As far as I can tell, AD has no intrinsic plausibility as a generalization from the finite. The manner in which the finite case of "determinacy" gets its plausibility seems to be clearly inapplicable to the infinite case. The finite case can be proved because one knows that a schema equivalent to it, is true, by quantifier logic:

$$(x_1)(\exists y_1)(x_2)(\exists y_2) \cdots (x_n)(\exists y_n) F(x_1, y_1, x_2, y_2, \dots, x_n, y_n) \vee \\ \vee (\exists x_1)(y_1)(\exists x_2)(y_2) \cdots (\exists x_n)(y_n) \neg F(x_1, y_1, x_2, y_2, \dots, x_n, y_n)$$

namely, one takes the negation sign in front of the matrix $F(x_1, y_1, \dots, x_n, y_n)$ in the second disjunct, and drives it to the front of the disjunct, getting

$$(x_1)(\exists y_1)(x_2)(\exists y_2) \cdots (x_n)(\exists y_n) F(x_1, y_1, x_2, y_2, \dots, x_n, y_n) \vee \\ \vee \neg (x_1)(\exists y_1)(x_2)(\exists y_2) \cdots (x_n)(\exists y_n) F(x_1, y_1, x_2, y_2, \dots, x_n, y_n)$$

which is of course true by excluded middle. But the Axiom of Determinacy would be equivalent to a schema containing an infinite number of quantifiers in front of a matrix:

$$(x_1)(\exists y_1) \cdots (x_n)(\exists y_n) \cdots F(x_1, y_1, \dots, x_n, y_n, \dots) \vee \\ \vee (\exists x_1)(y_1) \cdots (\exists x_n)(y_n) \cdots \neg F(x_1, y_1, \dots, x_n, y_n, \dots)$$

But how do we get an analogous result here? If we try to drive the negation out, where does it go? There is an infinite series of quantifiers in front of the matrix, and no latest place for the negation to land. The non-wellfoundedness of the series of quantifiers seems to destroy any credibility for AD that might derive from the finite case. The status of AD seems to be strictly one of a hypothesis entertained and worked with. I can contrive no intuition that would indicate $P(\omega)$ to conform to the axiom of determinacy. On the other hand, strong intuitions recommend the axiom of choice; we can easily conceive how the sets which verify the axiom of choice would arise.

I have said that the combinatorial notion of set is the most full-bodied form of mathematical platonism. It is in this light that one can perceive the common elements in its claims about the height of V , and about the breadth of V . In both cases, it holds V is as large as possible, is created in as arbitrary ways as possible. For inasmuch as this creation goes on out in mathematical reality, why should V not be just extraordinarily luxuriant? A consequence of this view of sets is that our knowledge of what V actually looks like is inconceivably impoverished. Perhaps it is this sort of view that led Gödel to the unusual position that the reflection principles should be accepted as intuitively obvious because of the unknowability of the absolute.

Some philosophers have taken the paucity of our knowledge of V to imply that we can never even refer to all sets. For, so the argument goes, our beliefs about sets are ineluctably so deficient that we can always construe our quantifiers to range over V_α for some α , and preserve all our beliefs. Why then not understand our quantifiers as ambiguous as to the height of the domain over which they range? Somehow, on this view, we do succeed in quantifying over all subsets of a set, if we quantify over the set itself, despite the fact that we may have very little idea of the details of how, say, the power set of a given ordinal is filled out. I find the difference between the ways the two cases are dealt with to be

unfounded. Let us return once more to measurable cardinals. Had we not conceived of such a cardinal, would this mean our quantifiers might be considered to range over a set V_k for k measurable, or rather over a kind of inner model of ZF in which a measure for this ordinal did not exist? This question does not appear to admit a determinate answer.

One can see this point another way. Recall the result that AD implies \aleph_1 is measurable. AD can be construed as entailing that $P(\omega)$ is extremely "thick". Suppose one for some reason were to doubt AC (as some do), and had not yet envisioned either AD or measurable cardinals, and, bizarrely enough, AD were true. Now surely we will have managed to refer to \aleph_1 . So if there were vagueness about what one's quantifiers were to range over, it would be better understood as vagueness with respect to how many subsets of ω they would range over, rather than w.r.t. the height of the universe. Of course, in the presence of choice it is possible to prove that if k is the first measurable cardinal, then there are k many inaccessible cardinals beneath it. And so if AC is true, as surely it is, a measurable cardinal is secured to be very high. But the result with AD may serve to suggest that exactly how high it is seen to be may depend critically on our assumptions about how the ranks of V should be fleshed out. And so the purported ambiguity of our discourse is not to be pegged exclusively on height.

Maybe some who believe there is this vagueness in our quantifiers are not easily troubled, and will not blanch at the possibility we cannot quantify even over all subsets of, say, ω . But it strikes me as more reasonable to surmise that we can so quantify, and that we can moreover quantify over all sets whatsoever. That I should not know how high V is or how wide it is, or more generally what are the truth values of the infinitely many formulas of ZF does not surprise me, nor does it make me any less inclined to think I am in ZF talking about all sets in V . After all, I do know this much about each set in V : the axioms of ZF hold for it. Why should not this knowledge avail in trying to talk about all sets? The incompleteness of my knowledge of V does not detract from my intention to be talking about all of V . This intention can be made so explicitly and emphatically that it would seem mildly perverse to go ahead anyway and construe my quantifiers over some set, instead of over all sets. For if sets really do exist independently of the mind, a belief the platonist holds so dear, how is it that this intention can miscarry? Jonathan Lear has argued⁶ that if we do not have the appropriate intentions toward a set we cannot quantify over it. For example, if we have never conceived of inaccessible cardinals, we cannot be said to have one in the range of our quantifiers. Now Lear never spells out what the suitable kinds of intentions are. While Parsons seems to have a view similar to Lear's in

this respect, he is no more explicit on this issue than Lear. And yet it is entirely critical to their claims that this be made clear.

Both Parsons and Lear presuppose that we can somehow succeed in quantifying over all subsets of ω , and that, in fact, if we can quantify over a set, we can quantify over all the subsets of that set. They need this first, weaker assumption to lend their view any credibility; failing the assumption, there is no ground for believing we are not quantifying only over a countable standard model of, say, all the true sentences of ZF. For there is only one way to assure that our quantifiers are not understood in this manner. We must be able to quantify over a subset of ω that codes up a function that collapses this countable standard model. The basic task facing a defense of Parsons' and Lear's view, then, is this. On what principled grounds can we say we can quantify over all subsets of (say) ω , which would not lead us to believe we can quantify as well over all sets whatsoever? Already I have urged that the connection between the height and the breadth of the universe is too intimate to make plausible the claim we can be confident we are quantifying over all sets in a rank, and be in doubt as to whether we are quantifying over all cardinals.

Let me further explicate some of these points. Return to the issue of intentions. In ZFT + choice it is easy enough

to prove there is a standard model for all the true sentences of ZFC. For example, if there is a measurable cardinal, there will be an ordinal in this model that satisfied 'x is a measurable cardinal'; and it will be true in this model that there are as many inaccessible cardinals, in the sense of the model, beneath this ordinal as there are members of the ordinal. All these splendid properties of this ordinal are cheaply got, however, by excluding any subsets of ω that would vitiate the structure. But what intentions are not captured by this model? All the truths of ZF are true here; hence in particular all the sentences we might ever believe in the language of ZF about the diverse ranks of ZF are true. So we are cut off from telling this story: we cannot be quantifying over all subsets of ω because there are sentences we might accept about $P(\omega)$ that are perhaps decided the wrong way in this model (certainly a possibility if we seek only a model for the theorems of ZF or some recursive extension thereof). Such a story would be of dubious value in any case, since a like story could be proffered on behalf of the view that we cannot take our quantifiers to range over V_k for some k. That is, we may in time adopt certain higher axioms of infinity, and yet they might be false if we take the range of the quantifiers to be some V_k that makes true all those axioms we accept now.

If there are principles that make it seem reasonable that we can quantify over all subsets of a set, but not over

all ordinals, they would seem to have to do with our somehow having a much firmer grasp of what an arbitrary subset of a set is, than we have of the autonomous, unconstrained generation of new ranks. Now it does impress me as correct that there is some metaphysical difference between the manner in which a subset of a set depends for its existence on the original set, and the manner in which the existence of a set depends on its members. Indeed, one might say that a subset of a set is in no way conditional on the set for its existence, but only on that subset's members. But I do not see how to parlay this ontic difference into a relevant epistemic one. If one has the idea that somehow we have to gather up the members of a set into that set, at each new rank, for these members to constitute a set, then it might be that there would be an open-endedness in how high we can quantify. But this is to see sets as at the core constructive, and the platonist will have no truck with that. However, I think it is this picture that operates in the back of one's mind when one thinks we cannot quantify all the way up. Once we divest ourselves of this picture, and acknowledge the independence of such set "generation" from our own minds, the openendedness in just one direction seems baseless.

There is but one other way to try to drive a wedge between the two different cases. And that is to observe that the cardinality of the subsets of a set, even taking into

account our use of replacement, will never furnish us all the iterations of new ranks that there are. And the number of such iterations is so stupendous that we cannot encompass them at once, while the number of subsets of a set, however arbitrarily engendered, we can. Here I can but invoke once more Cohen's suggestion to try to undermine this move. For it may be that there is no way we can attain the cardinality of the power set of a set working up by replacement from below. Yet we do think we understand what an arbitrary subset of a set is.

I would like to pursue now some further issues about the combinatorial notion. Parsons believes that Wang has captured the motivation for replacement when Wang says that:

Once we adopt the view point that we can in an idealized sense run through all members of a given set, the justification of SAR (i.e., replacement) is immediate. That is, if, for each element of the set, we put some other given object there, we are able to run through the resulting multitude as well. In this manner, we are justified in forming new sets by replacements. If, however, we do not have this idea of running through all members of a given set, the justification of the replacement axiom is more complex.⁷

Now I see the justification of the replacement axiom as arising more basically from the cardinality principle: if there are not too many sets in a multitude, they are bound up into a set. On my view, though perhaps not Cantor's, this basic picture would support not only replacement, but also power set (in conjunction with separation). In this latter case we know the set obtained is of greater cardinality; perhaps of

fantastically greater cardinality. In Cantor's 1899 letter to Dedekind, he proposed what was essentially the axiom of replacement, but curiously not the axiom of power set, as part of a new foundation for set theory. Now it may be this was just an oversight of Cantor's. On the other hand, it might have been in acknowledgement of the force of the power set-separation principle that he hesitated. For how could he know that in general, if one took a set, and then formed its power set, one did not go to a multitude so large as to be an "inconsistent multiplicity"? Or that the upshot of a repeated use of this principle was not such a multitude? It must have seemed quite evident to Cantor that the multitude of natural numbers was a set, and that arbitrary submultitudes of the natural numbers must be sets; for this much must, evidently, be true if analysis is to be at all possible. But analysis by itself does not require the existence of the set of all subsets of ω . Even Cantor's proof that there are more real numbers than there are rational numbers goes through unimpeded inside of classical analysis; here there is no obligation for the subsets of ω all to be fastened up into a set. So in postulating such a thing one must have the temerity to step beyond what classical mathematics would seem to uphold. In view of Kronecker's attack on Cantor, and Kronecker's claim that only the natural numbers really existed, it must have been difficult enough for Cantor to sustain even what analysis

seemed to demand.

However fair the foregoing may seem as a reconstruction of Cantor's views in 1899, it well represents a possible resting point in the development of the notion of set in ZF. The position is actually quite conservative, despite its embracing the axiom of replacement; for lacking the power set axiom, there is no way to engender sets of high cardinality.

What is all this in service of? Simply this. Parsons asserts Wang has, in the passage quoted, put well the intuitive underpinning of the axiom of replacement. But insofar as one sees the axiom as a principle whose chief purpose is to assure sethood by forestalling any explosion into an inconsistent multitude, a more fundamental picture would seem to underlie it. It is not that we, in some sense, take each element of the original set, replace it with its associated set, and then see that the resulting multitude is bound up into a set. The act of replacement does not figure importantly in our countenancing the multitude as a set; it is rather that we infer it is a set because we see it cannot have surged into an inconsistent multiplicity, since there is no surging at all. This is perhaps a rather subtle distinction; but it seems to adhere better to the underlying picture that justifies higher axioms of infinity, the axiom of choice, and separation. The following point may focus the distinction I have in mind. Our insight that a multitude is a set should, not the ideal case,

follow strictly why it is that that multitude is a set. Now the act of replacing each member of a set with that member's correlated set is not something that presumably transpires out in mathematical reality. Or, to remove the metaphor, the existence of the new set is not contingent in any way upon the existence of the old. As I have urged before, the only juncture at which it is clearly proper to speak of such contingency is when a set is said to depend for its existence upon its members. To repeat, replacement is best understood as capturing in part the idea that the iteration of levels occurs as often as we can conceive; this viewpoint makes it one with the higher axioms of infinity and separation in its deepest motivation.

In a footnote, Parsons criticizes Boolos (1971)⁸ for not having seen the ranks of V and the sets of V as being formed in a certain fashion together, so that if a well ordering were to come about in some rank in V , there should be a rank as high as that well ordering. But the motivation for replacement, as I have set it forth, would vindicate Boolos' original approach. For there is no sense in which the existence of a well ordering should oblige us to believe there is a rank as high, short of adopting the combinatorial notion, that justifies at once the full force of the replacement axiom. It is not as if such well orderings should be conceived to generate the new ranks, as Parsons' criticism apparently

presupposes. One thought that might seem, even so, to encourage taking well orderings to generate new ranks is this. If there were not such ranks, there would not be an ordinal for every well ordering, on von Neumann's identification of the ordinals. But this is a frail reed to rest so much upon, inasmuch as it is quite easy to identify the ordinals with certain equivalence classes of well orderings, using Scott's trick.

A parting remark on a very different issue. Parsons entertains the following possibility. Clearly we can construct a theory of ordinals that will not have the full power of ZF. Suppose then that we take all ordinals as individuals, rather than obtaining them in the usual fashion by set theoretic means. Then it seems appropriate to operate upon these ordinals as we would upon any multitude of individuals, and gather together in particular the set of all ordinals. But when this is done, it seems we can develop a version of the Burali Forti paradox, by defining a new relation \ll , which is a well ordering:

$X \ll Y$ iff x, y are ordinals which are individuals, and $x \ll y$ on that ordering relation, or y is the set of all such ordinals, and $x \in y$.

On this definition, the set of all ordinals will be greater in the sense of \ll than all ordinals; but \ll is a well ordering; hence the set of all ordinals would have to be ordinally greater than itself, since it would be an ordinal. Contradiction.

A certain observation will obviate this would be contradiction. We can in the typical fashion reduce ordinals to the sets of ZF. This means that ordinals just are sets; that is, to be committed to an ordinal is to be committed to sets. Of course, there are multifarious ways to identify ordinals with sets. But the fundamental point is that sets exhaust the mathematical universe (with the possible exception of categories); the reductions of the several branches of mathematics, including any theory of ordinals, to set theory should be taken to demonstrate just this fact. In consequence, it is no more legitimate to construe all ordinals as individuals from which sets can be formed than it would be to construe all sets as individuals from which new sets can be formed. And this latter we surely deem to be wrongheaded.

CHAPTER IV

The translation Parsons constructs between ZFT and NB, in the direction that concerns us, is this.¹ In ZFT, classes are understood as pairs (n,s) , n a formula of ZF, and s a sequence of sets. Take any occurrence of $\exists Y(\dots Y \dots)$, replace it by

$$\exists n \exists s (\dots \{x \mid \text{Sat}(n, s^0, x)\} \dots)$$

then eliminate the abstract. How convincing is this translation as a demonstration that classes just are such pairs? One problem with such an identification of classes with pairs is that, under the translation Parsons sets forth, for each class there will be more than one pair corresponding to it. Thus, the universal class will correspond to any pair of the form $(\ulcorner x=x \urcorner, s)$, regardless of the sequence s , on this translation. In this respect, the translation differs markedly from that of PA + ZF into ZF, since in the latter translation there is but one entity in ZF correlated with each number. It is true that there is a variety of ways to identify numbers with sets; we can use von Neumann's method or Zermelo's; but the relevant fact is that on a given translation there is a unique set correlated with each number. But if there is this failure in uniqueness, some question exists as to whether the translation provides an ontological reduction. What we might be seeking, in order to effect a reduction, is a proxy function, in the

sense of Quine.² That is, as Quine puts it:

We specify a function, not necessarily in the notation of θ or θ' [the reduced, and reducing theories, respectively] which admits as arguments all objects in the universe of θ and takes values in the universe θ' . This is the proxy function. Then to each n -place primitive predicate of θ , for each n , we effectively associate an open sentence of θ' in n free variables, in such a way that the predicate is fulfilled by an n -tuple of arguments of the proxy function always and only when the open sentence is fulfilled by the corresponding n -tuple values.³

Now is there a way of modifying Parsons' translation so that it does give a unique entity for each class? There is, but the most natural way of changing the translation presents some serious difficulties. Since Parsons' translation is one-many, the straightforward manner of altering it would be to pick out one of the many pairs related to each class as its proxy. However, as even the language hints, at least the axiom of choice is inextricably involved in this maneuver; indeed, as we shall see, an even more powerful principle is demanded. Now, to begin with, it would seem troubling for a principle like choice to be required for this sort of translation to be effected: intuitively, such a reduction would not appear to rest on such an axiom; if classes are sets, the axiom of choice should not be needed for us to see this is so.

Hence if it were NB+ plus AC which were being reduced to ZFT + AC, the reduction would already be suspect because of the presence of choice. However, as I have noted, to choose a unique pair for each class demands an even stronger background theory than ZFT + AC, namely what I shall call ZFT+ + WO. While we are not compelled to use this stronger theory, save if we wish to single out a pair among those that Parsons' translation furnishes, as the correlate of each class, we shall be instructed by examining this alternative.

ZFT+ + WO is the following theory. Take the language of ZF and extend it by adding a two place predicate, $R(x,y)$. Then WO is the principle:

$$\forall a (a \neq \emptyset \rightarrow (\exists! u) (u \in a \ \& \ R(a,u)))$$

To get ZFT+, enrich the truth theory by adding a clause for primitive predicate $R(x,y)$, supplying a truth theory for the extended language; finally, let separation be extended to include this new predicate, the strengthened Sat predicate, and first order compounds thereof.

Now we seek to reduce ZFT+ + WO to NB + AC+. Here AC+ is the following theory, closely analogous to WO:

$$\exists R \forall x (x \neq \emptyset \rightarrow (\exists! u) (u \in x \ \& \ \langle x, u \rangle \in R))$$

However, despite the strong formal similarities between AC+ and WO, we shall encounter reason to doubt that the plausibility of appending AC+ to NB should confer any merit upon the addition of WO to ZFT.

We want to find a proxy function between classes and pairs. As Quine points out, this must be done from a background theory that asserts that both exist. The background theory here, then, will be one that has at least the power of NB + AC+ + ZFT+ + WO. This is similar to the reduction of PA + ZF to ZF. For there the background theory is PA + ZF, which via a proxy function effects its own reduction to ZF. This latter reduction is accepted in the spirit of reductio ad absurdum, as Quine says; from a theory that embraces numbers and sets, we show that there is no need for anything but sets. The same motivation is present in our attempt to reduce NB + AC+ + ZFT+ + WO to ZFT+ + WO. In any case, the translation proceeds thus. In our background theory, NB + AC+ + ZFT+ + WO, we have the relation R such that for every nonempty set a, there is a unique u such that $R(a,u) \ \& \ u \in a$. Analogously to what Parsons shows, for every class X in NB + AC+, there will be some pair (n,s) such that

$$\forall x (\text{Sat}(n, s^0, x) \leftrightarrow x \in X)$$

where Sat is obtained from ZFT+. For each class X, take the set of all pairs of minimal rank which thus correspond to the class. Using relation R, choose for each class a member of the set so obtained for that class. That member will be the unique pair in ZFT+ + WO that corresponds to it. The rest of the translation goes through pretty much as before: we replace all occurrences of $(\exists x)(\dots x \dots)$ by

$(\exists x)[x \text{ codes up a class } \& (\dots \{z \mid \text{Sat}((x)_0, (x)_1^0, z)\} \dots)]^*$

Again, the abstract is eliminated. On this translation, the proxy function $F(X, z)$ would be defined thus:

$F(X, z) \stackrel{df}{\iff} \exists u [\forall y (y \in u \iff \forall z (z \in X \iff \text{Sat}((y)_0, (y)_1^0, z))) \& y \text{ is of min. rank } \& \& R(u, z)]$

Let us now review the circumstances that drove us to consider NB + AC+ and ZFT+ + WO in the first place. We were rightly committed to ZFT; but NB was translatable into ZFT. We speculated that to be committed to the entities of ZFT might be the same as being committed to the entities of NB, in view of the translation. However, it seemed that this final conclusion would not follow unless we could somehow reduce NB to ZFT, and this required a proxy function; such was not possible on Parsons' translation. In order to bring about the proxy function in a natural way, we adopted ZFT+ + WO. So, unless we are committed to ZFT+ + WO, and not merely to ZFT, we are not committed in virtue of the new translation to the existence of classes.

How credible, then, is ZFT+ + WO? Not credible at all, I am convinced. And the details of the defence of my answer are of considerable interest to us. I call WO by that name because it is equivalent to the existence of a definable well ordering of the universe of sets. WO may at first blush appear rather more appealing than $V=L$ (which in a sense

*Here 'x codes up a class' is the obvious expression.

implies WO, since $ZF + V=L + WO$ is a conservative extension of $ZF + V=L$, but in fact WO imposes a neatness on the universe that is, insofar as we treat it as a primitive predicate, even more difficult to believe. For the formula $R(x,y)$ that allots us our definable well ordering is not a formula of set theory, as is the formula ' $x \leq y$ ' that well orders the constructible universe; it is a primitive formula that we somehow manage to intend so that it well orders the universe. But how could we have such a nature, and the universe of sets have such a nature, that we could succeed at this?

Now a result of Easton's implies that $ZFT+ + WO$ is not equivalent to $ZFT +$ the axiom of choice for sets.⁴ But one might think that, despite the fact that $ZFT+ + WO$ is strictly stronger than $ZFT+ + AC$, the intuitions underpinning AC could be extended to justify an axiom like WO. After all, the WO principle seems to play a role in ZFT closely analogous to the role played by AC+ in NB, and powerful set theoretic intuitions support AC+ in the context of NB. Why may they not be taken to sustain WO in the context of ZFT? However, this line of argument would be misguided, I think, and it is of some importance to recognize why. Our set theoretic intuitions do indeed tend to uphold AC+ for NB; but these are intuitions about collections, not about extensions of predicates. We will allow there is a class that codes up a functional from each set to exactly one member of that set, but

this, I am persuaded, is because we see classes to be collections and therefore closed under many of the same operations sets are. Consider that when we ground the existence of such a class, we certainly do not go about it in the following fashion. First, we see that we can intend a predicate to be true of precisely each set and a unique member of that set; then we proceed to recognize the class that is the extension of that predicate.

Rather, our justification would advance thus. Once a collection is given, the subcollections that also must exist come about in pretty much arbitrary ways. (Of this matter I shall be writing at length in another portion of my thesis.) Now there is in NB a class that codes up the relation between each set and all the members of that set. Why should there not be a class that would be the subcollection of this coded relation got by restricting it to a function? This sort of justification appeals very strongly to what Paul Bernays had depicted as the combinatorial character of sets or collections.

As suggested above, it is difficult to see how one could somehow intend outright a primitive predicate so that it would be true of exactly each set and a unique member of that set; and it is difficult also to see how one could build up a complex predicate that would do the same job starting out with primitive predicates that are intuitively acceptable. In view

of this, one would anticipate that from the standpoint of ZF or ZFT, in which all collections are sets, and not proper classes, there would be no way to justify the use of a predicate as WO requires: the relevant combinatorial principles would not apply for a mere extension of a predicate. As I have observed earlier, predicates do have certain closure properties: for example, if there is a language meaningful to us containing a predicate with a certain extension, then there is a language meaningful to us containing a predicate whose extension is the complement of that extension. However it may be that we are able to do it, we are able to intend predicates so that the complementary predicate of a predicate has meaning if the predicate has it. In general, moreover, predicates are closed in this manner under the Boolean operations union, intersection, complement. These principles are quite distinct from the combinatorial principles applying to sets, since, for instance, we do not believe sets are closed under complementation: the complement of the empty set would be a set containing everything.

The observations above are relevant, though we do not have to use ZFT+ + WO to effect a translation, because they indicate a certain view on which the point of NB is not simply to supply certain entities, namely classes, as extensions of all the predicates of ZF. For if this were all that would ground NB, then NB + AC+ would not seem an intuitive extension

of NB, inasmuch as the motivation for AC+ is so different from what predicative considerations alone would allow. Parsons, in "Sets and Classes" argues that it is precisely as a way of giving any extension of a predicate a corresponding class that NB is adopted. This view, if correct, would seem to confer a certain credibility upon his conclusion that the quantifiers of set theory should be construed intuitionistically, by the following argument. Suppose the classes of NB are to be understood as merely the extensions of predicates of ZF. Now we are surely committed to ZF, and to the meaningfulness of the predicates of ZF, that is, to the fact that the extensions of the predicates of ZF are determinate. Since NB appears to be just a convenient way of capturing these commitments, we seem bound up with the classes of NB as objects. But the classes are set-like objects, and once we have allowed predicative classes, most of the same operations under which sets are closed would apply also to these classes. A natural next step would be to grant that the class asserted to exist by AC+ does exist. After that, there would seem to be little to prevent us from going further and countenancing impredicative classes, and, after that, little to stop us from taking all collections of classes to exist, etc. Intuitionistic-like quantifiers seem then to be forced upon us.

By now, it should be evident how to block such an argu-

ment. Either talk of the classes of NB is to be understood merely as a kind of shorthand for talk of the extensions of predicates in ZF, or talk of the classes of NB is to be interpreted as referring to set-like entities, with the combinatorial properties such entities must possess. In the first case, there is no reason to proceed from our acceptance of NB to an adoption of, say, NB + AC+. For that would be to treat the extensions as set-like objects, and by hypothesis we have avoided that. On the other hand, if the classes of NB are construed as set-like objects, then probably we want to advance to NB + AC+. We might want to go further to impredicative classes; but this move is not usually made, precisely because this would open up the iterative process once again on these classes. F. R. Drake, in Set Theory, puts the point this way:

This impredicative extension (Kelley-Morse Set Theory) has an unsatisfactory nature from the point of view of the cumulative type structure. If we consider V to be the universe of all sets, then classes are subcollections of things from V ; if we quantify over classes, this implies that we have the collection of all classes to talk about, and the collection of all classes would be exactly the thing we should take as the next level,

following all the levels used to make up V. In other words, talking about all classes is tantamount to saying that we have not taken all levels, with no end, but we have another one, the level of classes, which we have not used for making sets. From this point of view, it is more natural to regard classes as not forming a completed collection, so that we should not quantify over classes.⁵

Now we should take with a grain of salt Drake's claim that we are in Kelley Morse set theory talking about the collection of all classes; for his purposes, it suffices that we must be talking about all (with a very broad sweep of the hands) classes in Kelley Morse set theory. At all events, since any intuition that would uphold AC+ would evidently uphold Kelley Morse set theory, and since Kelley Morse set theory clearly does suggest further levels, it seems best to stop the regress at its root: the idea that the class quantifiers of NB should be construed objectively.

My own preference is to adopt NB, but interpret its class quantifiers in a semi-substitutional fashion, an alternative Parsons outlines at one point.⁶ This approach would seem to absolve us of any ontological commitment to proper classes in employing the theory of NB. Precisely because NB, with its class quantifiers construed objectually, is so

delicately poised on the brink of the excessive ontological commitment of Kelley Morse set theory, and because the intuitions that ground this delicate position are so tenuous, doing away with the extra ontological baggage of NB is for me the favored course. But naturally, if to be committed to ZFT is in itself to be committed to the ontology of NB, then this position is not tenable anyway; so let us return to this question.

How can a one-one translation directly between NB and ZFT be obtained? For each class, there is a set of pairs of minimal rank that corresponds to that class on Parsons' translation. Let this set be identified with the class. Whenever there is an occurrence of $\exists Y(\dots Y \dots)$, put in its place:

$\exists x [x \text{ codes up a class } \& \exists v \in x [\dots \{y \mid \text{Sat}((v)_0, (v)_1^{0,y})\} \dots]$

where again the abstract is purely virtual. 'x codes up a class' is defined: $x \text{ codes up a class} \stackrel{\text{df}}{\iff}$

$\stackrel{\text{df}}{\iff} \forall u, v \in x [\forall y (\text{Sat}((u)_0, (u)_1^{0,y}) \leftrightarrow \text{Sat}((v)_0, (v)_1^{0,y}))] \&$
 $\& \forall u, v \in x (\text{Rank}(u) = \text{Rank}(v)) \&$

$\& - \exists u, v \in x [(\text{Rank } v < \text{Rank } u) \& \forall y (\text{Sat}((u)_0, (u)_1^{0,y}) \leftrightarrow \text{Sat}((v)_0, (v)_1^{0,y}))]$
 The proxy function $F(X, y)$ is the formula:

$y \text{ codes up a class } \& \exists x \forall y \forall u (u \in X \leftrightarrow \text{Sat}(x_0, x_1^{0,u}))$

It should be noted that we cannot prove, in our present background theory, $(X) (E!y)F(X, y)$. But it is clear that this formula is true on the usual construal of the class quanti-

fiers in NB. If truth does not seem to suffice, we can always retreat to the metatheory for NB and prove it there. Now this last move may seem unsatisfactory, since it may seem to go beyond the spirit of reductio ad absurdum in which the reduction is to take place. Perhaps we can cheerfully say: from a theory with objects A we can see their superfluity. But our cheer may wane if we must draw on the force of its metatheory to prove the superfluity of objects A. Nonetheless, I think there is a redeeming feature in the present case. After all, we do have a proxy function, and a proxy function expressible in the original theory, NB. No danger of Pythagoreanism lurks if we adopt the above move for such cases. For example, while from the standpoint of ZFT one can prove the existence of a model in the numbers for all the truths of ZF, we do not thereby get a proxy function: cardinality considerations rule out this possibility. Moreover, no arithmetical predicate expresses the relation that is the interpretation of ' \in ' in this numeric model of ZF. But there is a predicate in ZFT (shortly to be defined) that expresses the relation that is the interpretation of ' \in ' (of NB) among the proxies for the classes of NB.

Let us suppose that the retreat to the metatheory described above is unproblematic. Can we then take the translation to effect an ontological reduction?

One first blush objection to the claim that there is a reduction might be put this way. There is a striking difference between the reduction of PA + ZF to ZF and that purported between NB and ZFT. PA + ZF is a two sorted theory in which there is no presumed overlap between ranges of the two kinds of quantifiers. But there is an intimate connection between the two kinds of objects, sets and classes; for all sets are classes; moreover, a proper class is understood to be distinct from any set--thus, for example, it is a theorem of NB that $\forall x(x \notin \{y \mid y = x\})$. Hence, $\exists Y \forall x(x \notin Y)$. In addition, the members of a proper class run all the way up the cumulative hierarchy; the proper class itself appears at no level in that hierarchy. What makes the reduction of numbers to sets entirely natural is just the failure of overlap between the laws that the two sorts of things, sets and numbers, must obey. Because we have no (or confused) intuitions about whether or not ' $3 \in 5$ ' is true, we have no serious misgivings when the entities identified with 3 and 5 bear or do not bear the \in relation to each other; likewise, we are not disturbed if $\emptyset + \{\emptyset\} = \{\{\emptyset\}\}$ or not.

It may seem to some that we do have intuitions about the truth of ' $3 \in 5$ '; namely, we can see it must be false; no number has any member. But I question whether the intuitions here appealed to have quite the character ascribed to them. I suspect that if we recoil at the suggestion that ' $3 \in 5$ ' is

true, it is because we do not see a number as the sort of thing that could have members; that is, our reaction to the claim that 3 is a member of 5 is that it is senseless, not that it is false. But, if this is so, then that very fact, paradoxically, should allow us to take ' $3 \in 5$ ' as true or false indifferently. For on the translation ϵ must hold or not hold between the proxies for 3 and 5, and since our intuitions are at base opposed to the sense of ' $3 \in 5$ ', we oppose those intuitions equally, whether we have ' $3 \in 5$ ' come out true or false.

In any case, we are not in a comparable situation with sets and classes. We do not want it to be that, somehow, the proxy in ZFT for the universal class of NB should be of lower rank in ZFT than the proxy for HF. And yet under the translation I have constructed it seems to be so.

It is not clear, however, what to make of this objection. True, the set corresponding to the universal class will be of lower rank in ZFT than the set corresponding to HF. But the set identified with the universal class will certainly not be of lower rank than that set in ZFT that corresponds to HF of NB, on the translation of the ϵ relation. This point may be more perspicuously put this way. Suppose x and y are sets in ZFT that code up classes. Then we say that x and y bear ϵ to each other on the translation ($x \epsilon_{NB} y$) iff

$$\exists u \exists z \in x [\forall v (v \in u \leftrightarrow \text{Sat}((z)_0, (z)_1^0, v)) \& \exists v \in y \text{Sat}((v)_0, (v)_1^0, u)]$$

On this translation of the \in relation, it is trivial that the entities of ZFT will bear \in_{NB} to each other when and only when they are so required by NB. Thus, the proxy for the universal class does not bear \in_{NB} to anything whatsoever in ZFT, and in that sense is not a "set". Just because the proxy function is a proxy function, we can expect to recover the new \in relation, \in_{NB} , to set things right among the entities correlated to the several classes.

It is not immediately obvious how effective this reply to the objection is. The reply seems to entail that no translation via proxy functions can have counterintuitive results of the kind delineated in the objection, since always (evidently), the various relations in the reduced theory can be recovered as the \in relation was, and our intuitions with respect to these recovered relations are of course exactly as they should be.

Let us consider another problematic case of reduction that follows in some respects the same general pattern of the alleged reduction of classes to sets, and see how the reply fits. We shall start with a theory that also has two sorts, where everything of the first sort is also of the second, but there is something of the second sort not of the first. Construct a theory exactly like ZF, save that it claims precisely one individual exists. This theory has two sorts, individuals and sets on the one hand, and individuals on the

other. Each sort has its distinctive variables, ranging over the obvious domains. Extensionality for sets will not be lost, since an axiom will be:

$$\forall x \forall z (-\exists a (x=a \vee z=a) \rightarrow \forall y (y \in x \leftrightarrow y \in z) \rightarrow x=z)$$

Where a is the variable for individuals, and x , y , and z range over both individuals and sets. We then take it as a further axiom that (a) $(\exists x)(x \in a \vee a = \emptyset)$. A proxy function from the ontology of this theory to that of ZF is pretty obvious. Send the unique individual a to the empty set; send the empty set to the singleton of the empty set. Now suppose all the members of a set x in this new theory have been assigned a set in ZF: then define the proxy of x to be the unique set containing all the proxies of its members; that is, the proxy functional $F(x,y)$ is:

$$F(x,y) \leftrightarrow (\exists f)(f: TC(x) \xrightarrow{int} \bigvee_{|rank(x)+2} \& f(a) = \emptyset \& f(\emptyset) = \{\emptyset\} \& \\ \& \forall z (z \neq a \& z \neq \emptyset \rightarrow f(z) = \{f(u) \mid u \in z\}) \& f(x) = y)$$

We say that $x \in_a y$ (where x and y are proxies) iff $y \neq \{\emptyset\}$ & $x \in y$.

Is this a reduction? Of course, the rejoinder to the objection would imply that it is; indeed the suitable \in_a relation is exhibited. I find myself somewhat reluctant to say that it is not. Now, one's first reaction may be to say that since the theory with the individual assumes there is an individual and ZF does not, and since individuals differ so in their properties from sets (individuals do not have unions with sets, after all!), the translation does not make for a

reduction. But, against this, imagine the theory of PA + ZF were altered in just one particular, namely, $\forall n \neg \exists x (x \in n) \& \forall n (n \neq \emptyset)$ were taken as an axiom (and therefore, of course, taken as well formed, unlike before). Would the addition of this axiom preclude the reduction of PA + ZF to ZF? Although von Neumann's and Zermelo's identifications of numbers with sets would no longer work, many straightforward identifications would; and I am inclined to the view that we would get a reduction under these identifications. A way of construing these identifications is again with a mind toward reductio ad absurdum: Assuming numbers are distinct from sets, even as axiomatic in our to-be-reduced theory, we can show they need not be taken as distinct. Similarly, we might understand $\neg \exists x (x \in a) \& a \neq \emptyset$ in ZF plus the individual, and in some such way $\exists Y \forall x (Y \neq x)$ in NB.

Perhaps at this point the discussion of whether there is an ontological reduction between NB and ZFT is becoming a bit too diffuse. We seem to be clutching at intuitions that would go one way or the other on the particular cases. Some reflection on more general considerations seems in order.

When it is in mathematics that one sort of object is reduced to another must be a vexing question, since the notion of mathematical object is vexing. Our connection to any mathematical objects, if they indeed exist, is decidedly tenuous; but to determine that a particular object, say the

universal class, can or cannot be viewed as identical to the ordered pair $\langle 'x=x', s \rangle$ may seem to rest too much on this already frail link. It is in any event quite obvious that our conception of a successful mathematical reduction and our notion of a mathematical object are so bound up that we must deal with them both to deal with either. We shall start with some thoughts about mathematical objects.

Naively, the most appealing view on the ontological status of mathematical objects is, perhaps, a platonism of the sort Godel espoused. Godel's position is that mathematical objects exist independently of us and constitute a well determined totality. As with physical objects, what mathematical objects there are is determined by the nature of the external world. Every sentence in the language of mathematics has a determinate truth value, whether we can determine it or not, if we intend our sentences in the natural way, so that universal quantifiers range over all collections (i.e., all mathematical objects), and ' \in ' is interpreted as membership. Indeed, it may seem difficult to separate Godel's view that the world of mathematical objects forms a well determined totality from this view that we can so intend our quantifiers and our primitive predicate(s). For it may seem constitutive of the notion of a well determined totality that we are able to employ such intentions; if we cannot in principle find out the truth value of all the sentences we

use, we can accommodate this by thumping hard on the distinction between ontology and epistemology; but if we cannot intend our quantifiers to range over all objects, or our primitive predicates each to have a unique interpretation, does it make sense to speak of the world out there as being nonetheless entirely well determined? In a few pages, we will consider this question at length.

In any case, there will be more on this in due time. Now let us observe that if we can employ our quantifiers and primitive predicates as Godel thinks we can, then the truth values of all our mathematical sentences are determined. This seems trivial enough to see, by induction. Suppose that φ is of the form $\psi_1 \vee \psi_2$ or $\neg\psi$. Then if it is determinate whether s satisfies ψ_1 and ψ_2 and ψ , then so is it whether s satisfies φ : it is not the truth functions that make for possible indeterminacy of truth value (or, more precisely, satisfaction value). Now assume that it is determinate for any sequence s , whether or not it satisfies $\varphi(x)$. Then, inasmuch as we can so intend the universal quantifier that it ranges over all objects, the satisfaction value of $\forall x \varphi(x)$ must be fixed by that intention. For this, evidently, is what it would mean to succeed in referring to all objects. Indeed, in the usual case, when philosophers say that a certain mathematical sentence is indeterminate, it appears to be the interpretation of the quantifier that they question. For example, one often

hears the continuum hypothesis is indeterminate for the following reason. How do we know that for any set of subsets of ω , there will exist a one-one correspondence between that set and either ω or the set of all subsets of ω ?⁶ For what, proceeds the objection, is to be included under the 'any', the 'all', and the 'there exists'? In fact, routinely when the independence of the CH or AC or virtually any other assertion is proved, it is accomplished at least in part by changing the domain of the starting model, and therefore the interpretation of the quantifiers.

Finally, there are the atomic predicates. It might seem that the atomic predicates, surely, are unambiguous in their interpretation. For instance, given sets a and b , it would appear perfectly determined from what we mean by ' ϵ ' whether $a \epsilon b$ or not. However, there are an infinite number of relations involving sets that are completely isomorphic to the ϵ relation; that is, there is a one-one map F from V to V such that $x \epsilon y$ iff $F(x) \epsilon_* F(y)$. An example of such a map can be obtained by the following device. Define $X \epsilon_* y$ thus:

$$x \epsilon_* y \text{ iff } x = \langle u, 1 \rangle \ \& \ y = \langle w, 1 \rangle \ \in U \cup w. \quad \text{Let } F(x) = \langle x, 1 \rangle.$$

Then, clearly, this new ϵ_* relation is isomorphic to the old ϵ relation; hence, precisely those sentences will be true in this model interpreting ' ϵ ' as ϵ_* as would be interpreting

' ϵ ' as ϵ . Now are our intentions about how we mean ' ϵ ' so unequivocal that they pick out the "real" ϵ relation of V , and not the ϵ_* relation of the new interpretation? It is rather difficult to escape the conclusion that, at least to some extent, our intentions in using ' ϵ ' are vague, and that in particular they cannot discriminate between ϵ and ϵ_* . One might respond to this: But we do have the intuition that we know perfectly well what ' ϵ ' means, as we ordinarily use it; we mean member of, and clearly the ϵ_* relation is not that relation! But is this sort of intuitive appeal anything more than a retreat to a background theory in which ' ϵ ' is already understood? Naturally, from the standpoint of a background theory in which we have already at hand a certain ϵ relation, and concomitantly a certain universe V , the distinction between ϵ and ϵ_* is quite straightforward. But the problem is that our use of the background theory might not be such that the ' ϵ ' predicate in that theory must be interpreted as ϵ and not ϵ_* .

With regard to physical objects, one's intentions seem to suffer less from this sort of vagueness. In that realm, the palpability of the objects, which allows such devices as ostension, seems to fix intended interpretations rather more decisively. If we want to rule out as unintended certain interpretations of some predicate that holds between macroscopic objects, we can very often point to some objects that

would be related under the unintended interpretation, and then deny that they are in the extension of the predicate as we wish to construe it. Ostension of this kind is evidently not available to us for mathematical objects. At best, we can employ what Quine calls deferred ostension, e.g., pointing to the symbol ' ϕ ' and meaning to refer to ϕ . But deferred ostension is of little help here, inasmuch as the difficulty then becomes to show we have managed to refer to one mathematical object and not another when we point to ' ϕ '. To use another example, suppose we wished to rule out the ϵ_* relation as inappropriate as an interpretation of ' ϵ ', in a manner like that described for predicates of macroscopic objects. We might point to ' $\langle \phi, 1 \rangle$ ' and then to ' $\langle \{\phi\}, 1 \rangle$!' and finally deny the two entities thus referred to by deferred ostension bear the ϵ relation to each other. But what is it we have referred to by pointing as ' $\langle \phi, 1 \rangle$ ' and ' $\langle \{\phi\}, 1 \rangle$ '? Are they $\langle \phi, 1 \rangle$ and $\langle \{\phi\}, 1 \rangle$ or rather ϕ and $\{\phi\}$? If the latter, then we have done something we are concerned to avoid: ruling out the genuine ϵ relation as a possible intended interpretation of ' ϵ '. Now how it is, exactly, ostension might work well for macroscopic objects, and not so well for mathematical objects, is a problem of considerable subtlety; most often, such an account appeals to causal connections between us and physical objects. There are those, however, who have no truck with this or any other way of making a distinction between how we refer

to physical objects, and how we refer to mathematical objects: perhaps Putnam, in "Realism and Reason" and in "Models and Reality", can be read to support this position. I am persuaded that there is an important difference between the two cases; what this difference is should soon become evident. But first we must discuss what mathematical objects might be, and why they are both problematic and inevitable.

The mathematical objects picture that to Godel is flesh of his flesh and bone of his bone is one difficult to assimilate for others. Those of us who have said in our hearts there is no transcendent reality will not sit comfortably with this view; for the differences between mathematical and transcendent objects are less profound than their similarities; both transcendent and mathematical objects are not locatable in space and time; both are causally inert; both are eternal. And yet there are not in the field many credible alternatives to a belief in mathematical objects. Kreisel has said that for a philosopher of mathematics what is at issue is not the existence of mathematical objects, but the existence of mathematical objectivity.⁷ Perhaps. But here we are, stuck, it seems, with theories that say there is a set of all numbers, there is an uncountable cardinal, and more embarrassingly extravagant things even than these. And it is difficult to make out how we are going to avoid taking these sentences to mean exactly what they appear to say. Indeed, the only view

that appears to get round the mathematical objects picture is the one proposed by Putnam, and lately taken up by Parsons, on which the notion of possibility assumes the burden of supporting mathematical objectivity. Insofar as we have serious doubts about the notion of possibility, this move avails us nothing, however. Intuitionism, which might seem to offer succor in this extremity, seems in fact no less otherworldly in its commitments than platonism. For intuitionism is scarcely a species of finitism: the varieties of mental constructions that must be real for intuitionism to be plausible already far outstrips any mental constructions we actually have or ever will have. On the score of remoteness from the everyday world, intuitionism seems no better off than platonism.

Evidently, some think Quine has a kind of platonism that avoids the problems with transcendence troubling Godel's more "naive" view. I do not see this. Quine's view is that we need to posit mathematical entities in order to do physics, and we need to do physics in order to explain sensory stimulations. So it is only the program of accounting for such stimulations that leads us to posit mathematical entities. This supposed fact is purported to make empirically respectable the existence of mathematical objects. But it seems really not to bear on the issue at all. For we surely feel an obligation to explain how it is we can know anything about the physical objects we posit, over and above observing we must

assume them in any explanation of sensory experience. That is, the task of explaining how we come to know about physical objects is one undertaken chiefly by physiologists, neurologists, and neuropsychologists. The task of accounting for our sensory experience is more in the province of the physicist, the chemist, the biologist, etc. Insofar as these two tasks are separate, we will not discharge our obligation to explain how we come to know about physical objects by accounting for our sensory experience. It may be said that the total science that encompasses both these enterprises manages to fuse them into one. But certainly the orientations of the two endeavors are quite distinct: one starts with sensory stimulations as given, and proceeds to posit various objects to give these stimulations coherence and intelligibility; the other starts with the physical objects as given, and attempts to show how they interact with our sensory organs, and how our sensory organs interact with our nervous system to furnish us with just those sensory experiences we have. And so long as the orientations are different in this way, we may ask: How is it we are exempt from an obligation to show how, assuming mathematical objects as given, we can come to know about them? This question Quine's view does not address.

There is a certain historical irony in the attention nowadays paid to the problem of how we can come to know about mathematical objects. For this problem is the inverse problem

of one Descartes dealt with. Descartes was at great pains to explain how it is that we, as immaterial substances, could manage to perceive the material world. In our more naturalistic age, we are instead perplexed by how we, as material objects, can have knowledge of immaterial mathematical objects. Descartes dealt with his problem by locating the scene of interaction between the immaterial mind and the material world in the pineal gland. Doubtless some philosopher, alert to the relevance of bygone philosophy to contemporary thought, will find in the pineal gland a solution to our problems, mutatis mutandis.

At all events, Mark Steiner has a different way out of our present difficulty;⁸ in my estimation, no more viable. His idea is in brief this. To give a causal account of how we acquire knowledge of something, we must appeal to some background theory. But "the axioms of analysis, as interpreted by the platonist, will indeed necessarily be used in whatever causal explanation can be given of our belief that the axioms, again as interpreted by the platonist, are true." This somehow absolves us, on Steiner's view, of providing further explanation. I fail to understand this. Presumably, any background theory that we would employ to give a causal account of our knowledge of physical objects would assume their existence. But does the presence of such an assumption

in itself qualify the theory as an explanation of this knowledge? No: we demand more of the theory; the theory must show how it is that physical objects interact with us. Our current puzzle is precisely this: What account of our interaction with mathematical objects can be obtained, or why should no such account be sought? To require a causal account of our "interaction" with mathematical objects seems out of the question; even to demand an account of our "interaction" is to ask too much. But what is left?

We see, then, some of the difficulties accompanying the view that we can know about mathematical objects. These difficulties in epistemology may derive from difficulties in the philosophy of language. Lately, many philosophers have grounded the claim that knowing about an object requires some special causal connection to it, in the more basic claim that even to refer to an object requires such a connection. The plausibility of the causal theory of reference places a great onus on the platonist. He must explain how we can refer to mathematical objects, and how we can indeed intend our predicates so that they pick out appropriate relations. As I have noted, Godel had the somewhat quaint view that we can employ the language of mathematics, i.e., the language of set theory, so that it picks out a unique intended interpretation. But Godel himself did not seem to take this view with great seriousness at all times. When arguing for the determinateness of the truth value

of CH, Gödel feels compelled to say we can develop intuitions about axioms that would decide CH, and to say it is in part the possibility of such intuitions that gives the question of the truth value of CH meaning. At no point does Gödel propose an argument to show that, for any sentence of set theory, such intuitions may be forthcoming. If, as Gödel claims, the term ' ϵ ' picks out but one relation, the membership relation, why do we need such intuitions to give the question of the truth value of CH meaning? In fact, if ' ϵ ' picks out just one relation, it seems clear it is not the possibility of such intuitions that serves to do this. For suppose all the sentences of set theory were determined by such intuitions; there would yet be an infinite number of relations ' ϵ ' could be interpreted to mean, with all true sentences coming out true on each such interpretation. If we can so employ ' ϵ ' that it latches onto a unique relation, that we can do so must not be only because we determine the truth values of the sentences of set theory. Indeed, it seems Gödel makes a move subversive to his view when he looks to intuitive axioms to fix the truth value of CH. For he thereby implies our acceptance of theories plays the primary role in effecting the interpretation of our predicates. But, evidently, if we are to secure a unique interpretation for the ' ϵ ' predicate, we must conceive the enterprise of interpreting language to be largely independent of what theories,

in detail, we adopt.

I say "in detail" here, because it may be that the enterprise of interpreting language uniquely cannot get moving, unless we hold some appropriate theory. Thus, it may be if we do not hold ZF, or some significant subset of ZF, we do not have good enough a grasp of ' ϵ ' to give it any interpretation, much less a unique one. But, certainly, if we are to obtain a unique interpretation for ' ϵ ', we must at some time go beyond the consideration of theories we do or might adopt.

Godel has more to say about the determinateness of the truth value of CH. If an axiom is fruitful, Godel claims, we have a prima facie obligation to accept it; and, since such an axiom might decide CH, the truth value of CH is grounded in this fashion as well. This view is particularly vexing. There is the problem just mentioned, the implicit assumption that we can fix the interpretation of language only via the theories, in detail, we accept. Again, for Godel, this assumption has intolerable consequences. But aside from this problem, there is another. Why should we take the fruitfulness of an axiom as evidence for its truth? This is a serious difficulty, because there might well be a pair of sentences with the following feature: one sentence in the pair decides (fruitfully) certain open questions in one direction; the other sentence decides (fruitfully) the very same questions in

the opposite direction; and both sentences are consistent with everything we believe. I can descry no reason to rule out such a pair. Indeed, if one is unsure of AC, AD and $V=L$ will constitute such a pair; for AD requires AC to be false, and $V=L$ requires AC to be true.

Let us now draw a certain inference from our discussion of Godel. It is highly dubious that considerations of fruitfulness (even if allowed) combined with considerations of intuitive evidence, could ever decide the infinite number of sentences of set theory. We are only finite beings, and the human race probably will not survive forever. And, should humans not as a race always be around, it seems obvious that only a recursive set of axioms will be intuited, or be found fruitful. But even if we did continue on endlessly, we would be little better off, at least in an endeavor to secure the truth values of all mathematical sentences. For when one comes to appreciate the general applicability of diagonal arguments, one begins to suspect the worst: There are sentences of set theory that, 1) do not follow from anything we, as finite beings, can manage to intuit, because they are so complex, and, 2) will not be fruitful for any questions less complex than themselves. But the existence of such sentences would, even if we should live forever, make it impossible to determine the truth value of each sentence of set theory at some time.

From the foregoing, we may conclude that none of the set theories to which we do, or even might assent are complete. But suppose we can pick out a more select group of interpretations than that consisting of those interpretations compatible with the theories we do or might accept. It cannot be that we pick out some group of interpretations consisting of all interpretations compatible with some complete theory; for there is no complete theory involved in our attempt to fix an interpretation. Rather, our attempt, if successful, must be a direct matching up of a predicate and a relation. Once it is granted that there is no way of achieving this reference to a unique relation, we seem stuck with an unappealing view. Namely, we cannot rule out a relation as an interpretation of a predicate if such an interpretation is compatible with those theories we might ever accept.

Now let us see the relevance of the foregoing to the issue of ontological reduction. There is at first blush, and I believe on later blushes, a connection between two questions. Suppose we are committed to a certain theory. Consider the ontology of this theory. A number of maps will exist that go from this ontology into or onto itself. A great many of these maps (an infinite number in general for theories with infinite ontologies) will preserve the truth values of the sentences of the theory under some translation of the predicates; of the maps that preserve truth, many will preserve further

important properties. Now one might take a select portion of these maps to show that our commitment to the ontology of the theory is no different whether we embrace the original structure, or the structure we get by the map; this despite the fact that the new structure may look very different from the original structure, from the point of view of the original structure. Which maps can occasion this sort of "indeterminacy" is the first question. Now Quine allows that any one-one map expressible in the theory will count as such a map. I will argue that this is too generous.

In any event, the second question has to do with ontological reduction. When one theory is reduced to another, the structure of the reduced theory is mapped via some functional into the reducing theory. Typically, this functional can be expressed in the reducing theory. Given two theories, and the ontologies of these theories, there may be any number of maps between these two ontologies that preserve the truth of the sentences in the first theory, under some translation of the predicates. Which of these maps are to count as providing ontological reductions? This is the second question.

Here Quine again allows that any map expressible in the reducing theory will effect a reduction. Quine's consistency in his treatment of the two cases I admire, and is precisely to my purpose; I disagree only about the promiscuousness of his constraints in both cases. Shortly I shall pre-

sent some further constraints for both cases; constraints that must come to play precisely because of such cases as the failure of the would-be reduction of NB to ZFT. For now, let us attend to the unity of these two questions.

To take the two questions to be ultimately one is to reinforce the view of ontological reduction I set forth in the first chapter. There I urged that to show one ontology can be reduced to another is to show that commitment to the reducing ontology requires commitment to the reduced ontology. For consider how we would justify seeing these two questions as one; presumably, it would proceed like this. 1) Cases of ontological reduction just are cases in which commitment to the reducing ontology requires commitment to the reduced ontology. 2) The case of the ontology of a theory being mapped into itself, in which commitment to the structure obtained by the map is the same as commitment to the original structure, is but a degenerate case of 1); that is, when a theory's ontology is mapped into itself, the theory is its own reducing theory, and to be committed to the ontology of this theory is to be committed to the ontology of the theory that has been mapped into itself, namely, its own ontology. Perhaps an example will be of some service here. Suppose we show in Peano Arithmetic that all even numbers can, on the appropriate map, and translation, model PA. Then, considering the reducing theory to be PA, we can say that to be committed

to its ontology requires commitment to the ontology of the theory modelled in it; this theory is, of course, PA itself. That is, from the standpoint of PA, one can see that commitment to its even numbers is the same as commitment to all its numbers; this is what an ontological reduction of numbers to even numbers would mean.

As an aside, let me observe that, with the preceding in mind, we can obviate a certain argument of Leslie Tharp. He writes:

It is rather startling to reflect that there are many self-reducing theories, and in fact the most important theories have this property (the property of being able to model themselves in a proper subdomain). For example, the set of numbers larger than 16 can be proved in arithmetic to form a domain of a model of arithmetic. A similar situation results in set theory if one lets $\{\emptyset\}$ take the role of \emptyset and extends the correspondence in the obvious way; here the same relation is used in the submodel as in the starting model. No one could maintain that these are interesting examples of reduction; so they are instructive in that they tend to illustrate what reduction cannot be about.⁹

Now I do not know if I can maintain that Tharp's examples are generally interesting cases of reduction; I find them interesting; but more to the point, I find them to be cases of reduction.

Now let us observe that, in cases of ontological reduction, we seem to require something quite strong: all true sentences in the language of the reduced theory must be true on the furnished translation into the language of the reducing theory. This might seem to leave us in an odd situation. On the one hand, we have granted that there is a sentence of set theory such that our use of ' \mathfrak{C} ' is indifferent to whether the interpretation of ' \mathfrak{C} ' makes that sentence true, or whether it makes it false; let us say that CH is such a sentence. On the other hand, we are presently claiming that if the ontology of ZF were to be reduced to another, all true sentences in the language of ZF must be true on the provided translation into the language of the reducing theory, including CH, or the negation of CH, whichever (unknown to us) is true. But this situation may not be as unnatural as it may appear. Ontological reductions hold, at the base, between structures, ontologies; the indeterminacy that may exist in the interpretation of our language arises from a relation between words and structures. The strictures for ontic reduction we might thus expect to be more severe than those for the fixing of interpretations of our language;

for these two would seem to differ in much the same way as what there is differs from what we can know.

I think that more than truth must be preserved in any map and translation that would provide an ontological reduction; indeed, I think the preservation of truth derives from the fact that more basic, algebraic features must be preserved. But let us return to the cases of clearcut, and possible reductions we had considered earlier. One salient feature of the reduction of PA + ZF to ZF is the naturalness of the transition from the ontology of the one to the ontology of the other. The manner in which the number theory of PA + ZF is embedded into the \in structure of ZF is quite nice algebraically, in the way that the \in structure reflects the operations of successor, addition, and multiplication. Paul Benacerraf, in "What Numbers Could Not Be", argues that if a progression is to model the natural numbers, that progression must be recursive.¹⁰ This constraint strikes me as reasonable, although perhaps requiring some reformulation; I believe that the progression must be computable, because the structure that is embedded into ZF must derive its features in a direct way from the structure of ZF itself. Thus the progression of numbers is certainly recursive from the standpoint of PA; but then the structure that is embedded in ZF, which is to model PA, must derive its recursiveness straightforwardly from computable relations in ZF.

Now the notion of recursiveness is ordinarily defined using arithmetical predicates. Hence it may seem wrongheaded to insist that the relation between the proxies of each number and its successor be recursive; that notion could only make sense, presumably, after the translation has been effected. That is, a relation is recursive depending on whether it can be expressed by certain simple arithmetical predicates; the progression of proxies for numbers, whatever that progression may be, is clearly going to be one of these, under the translation of the arithmetical predicates.

I think the nerve of Benacerraf's point, however, is unaffected by the foregoing consideration. The basic intuition lying behind Benacerraf's point is, I think, this. It is appropriate to identify the progression of numbers with, say, the progression of the ordered pairs that standardly code up formulas. But it is not appropriate to identify the progression of numbers with the subprogression of the arithmetical truths, since this new progression is too complex in its nature. Now this complexity can be adequately characterized strictly from the standpoint of set theoretic notions, without introducing the notion of recursiveness. For the basic notion of computable function can be understood from this standpoint; indeed, the notion of computable function is par excellence a notion that seems to be amenable to formulation in many ways, with recursive function being just one such

formulation. The notion of computable function, in its perhaps most intuitive characterization, is cast in terms of finiteness, and the notion of finite can be captured in set theory. The characterization I have in mind is this: A function is computable if it takes finite objects to finite objects using an algorithm, which is of finite length. It is thus quite easy to formulate the notion of a computable function using strictly the language of set theory. And it is, of course, quite easy to show that the progression of arithmetic truths is not computable. So from set theory we can see that some progressions are less complex than others. We recognize that the relation between a number and its successor should be as simple as possible; hence we insist that the relation be a computable one.

Perhaps another example will help convey the point about how structures should be embedded to effect reductions. I have presented what I take to be a reduction of ZF to ZF, using the map $x \mapsto (x, 1)$. Why do I consider this a reduction? Consider the new relation that interprets ' \in ' under the translation; that is, the relation \in_* such that

$$x \in_* y \leftrightarrow \exists u \exists v (x = \langle u, 1 \rangle \ \& \ y = \langle v, 1 \rangle \ \& \ u \in v)$$

This new \in relation must derive its important mathematical properties from like properties of the relations of ZF. And this, of course, it does; for example, the well-foundedness of \in_* comes about in a direct manner from the well-foundedness of \in .

Contrast, now, the case of NB and ZFT. One sees how unnaturally the \mathfrak{E} structure of NB is reflected in ZF, under the map and translation furnished, in the artificiality of the definition of \mathfrak{E}_{NB} : there is virtually no important relation between those things related by \mathfrak{E} under ZFT and those related by \mathfrak{E}_{NB} . For the \mathfrak{E}_{NB} relation involves the Sat predicate in its definition, and indeed in any translation between the two theories the Sat predicate or the like would be involved essentially. Symptomatic of the radical rearrangement of entities under \mathfrak{E}_{NB} is the fact that the proxy in ZFT for the universal class of NB is, under \mathfrak{E} in ZFT, only finitely high; this despite the fact that the universal class is as high, in the \mathfrak{E} structure of NB, as is possible. Now perhaps if this were the only anomaly, we would still be willing to call it a reduction; but such anomalies are systematic and inescapable, and I am persuaded that we are loath, for this reason, to think NB can be reduced to ZFT.

In addition to the foregoing considerations, which are rather subtle, and clearly not yet fully developed, there are more obvious considerations bearing on the establishing of ontological reduction. Certainly, one necessary condition of ontological reduction would be this: The reduced structure and the structure that mirrors it in the reducing structure must have the same cardinality. Or at least this is so if we do not have demonstrably superfluous entities running about,

as in the case of certain structures that Richard Grandy has contrived. Such cases are easy enough to mark off from the rest that they can safely be considered a separate species, as Quine argues.

In any event, the proxy function requirement Quine imposes, and which we have considered, may be seen as, in part, a way of preserving cardinality from the reduced structure to its mirroring structure. But there is more that can be said for it. It figures, in an obviously crucial fashion, in providing this mirroring structure, and in guaranteeing that it will mirror in the strongest possible manner; that is, that it will be isomorphic to the reduced structure.

Chapter I

1. 'Sets and Classes,' Nous 8, March 1974, pp. 1-12;
'In formal Axiomatization, Formalization, and the Concept of Truth,' Synthese 27, May-June 1974, pp. 27-47; 'The Liar Paradox,' Journal of Philosophical Logic 3, pp. 381-412, Oct. 1974.
2. 'Outline of a Theory of Truth,' Journal of Philosophy 72, November 6, 1975, pp. 690-716.
3. 'The Concept of Truth in Formalized Languages,' in Logic, Semantics, Metamathematics, Oxford" Clarendon 1956.
4. Hans Herzberger discusses this in 'Paradoxes of Grounding in Semantics,' Journal of Philosophy 67, March 26, 1970, pp. 145-166.
5. 'Sets and Semantics,' Journal of Philosophy 74, Feb. 1977, pp. 86-102.
6. In the papers previously cited.
7. 'Informal Axiomatization, Formalization, and Truth,' p. 66.
8. Ontological Relativity, New York: Columbia University Press, 1969; pp. 59-60.
9. This view of Frege's is discussed in Paul Benacerraff's 'What Numbers Could Not Be,' Philosophical Review 74 (1965), pp. 47-73.

Chapter II

1. 'Sets and Classes,' pp. 4-6.
2. See for example, 'God, the Devil, and Grödel,' Monist 51, January 1967, pp. 9-32.

Chapter III

1. 'On Platonism in Mathematics,' in Paul Benaceraff and Hilary Putnam (eds.) Philosophy of Mathematics: Selected Readings, Englewood Cliffs, NJ: Prentice-Hall, 1964, p. 276.
2. In From Frege to Gödel, ed. by Jean van Heijenoort, Cambridge: Harvard University Press, 1967; pp. 113-118.
3. 'On Platonism in Mathematics,' p. 276.
4. Set Theory and the Continuum Hypothesis, Reading, Mass: W. A. Benjamin, Inc., p. 251.
5. For a summary of these results, see E. Kleinberg, Infinitary Combinatorics and the Axiom of Determinateness, Berlin: Springer-Verlag.
6. 'Sets and Semantics.'
7. From Mathematics to Philosophy, London: Routledge and Kegan Paul, 1974, p. 184.
8. 'The Iterative Conception of Set," Journal of Philosophy 68, pp. 215-230, April 22, 1971.

Chapter IV

1. 'Sets and Classes,' pp. 4-6.
2. Quine discusses this in 'Ontological Relativity' (in Ontological Relativity) and in 'Ontological Reduction and the World of Numbers' (in Ways of Paradox, Cambridge: Harvard University Press, 1976).
3. 'Ontological Reduction and the World of Numbers,' p. 218.
4. Powers of Regular Cardinals, Thesis, Princeton University.
5. Set theory, New York: Elsevier, 1974, p. 17.
6. 'A Plea for Substitutional Quantification,' Journal of Philosophy 68, 1971, p. 235.
7. This ascription to Krisel is found, among other places, in Putnam, Mathematics, Matter, and Method. London: Cambridge University Press, 1975.
8. 'Platonism and the Causal Theory of Knowledge,' Journal of Philosophy 70, Feb. 8, 1973, p. 61.
9. 'Ontological Reduction,' Journal of Philosophy 68, March 25, 1971, p. 157.
10. 'What Numbers Could Not Be.'