The set-theoretical program of Quine succeeded, but nobody noticed

M. Randall Holmes

Boise State University, Boise ID 83725 U.S.A.

1 Abstract

The set theory "New Foundations" or *NF* introduced by W. V. O. Quine in 1935 is discussed, along with related systems. It is argued that, in spite of the fact that the consistency of *NF* remains an open question, the relative consistency results for *NFU* obtained by R. B. Jensen in 1969 demonstrate that Quine's general approach can be used successfully. The development of basic mathematical concepts in a version of *NFU* is outlined. The interpretation of set theories of the usual type in extensions of *NFU* is discussed. The problems with *NF* itself are discussed, and other fragments of *NF* known to be consistent are briefly introduced. The relative merits of Quine-style and Zermelo-style set theories are considered from a philosophical standpoint. Finally, systems of untyped combinatory logic or λ-calculus related to *NF* and its fragments are introduced, and their relation to an abstract model of programming is outlined.

Contents

2 Introduction

The set-theoretical program of Quine to which we refer began with the system now called *NF* introduced in his article "New Foundations for Mathematical Logic" ([24]) in 1935. *NF* is appealing because it is very simple. It is the first-order theory with equality and membership with the axiom of extensionality (objects with the same elements are equal) and the axiom of "stratified" comprehension ($\{x \mid \phi\}$ exists for each stratified formula ϕ). A formula ϕ in the language of *NF* is "stratified" if and only if it is possible to assign a non-negative integer type to each variable occurring in it in such a way as to obtain a well-formed formula of the simple theory of types. *NF* can be finitely axiomatized using a short list of particular instances of stratified comprehension (this was first established in [11]).

It is fairly well-known that the problem of the consistency of *NF* is very hard; it remains open after 50 years, and we do not claim to have solved it. Specker proved in 1953 (in [30]) that the Axiom of Choice is false in *NF*, which discouraged mathematicians from working with this system and raised doubts as to its consistency. Nonetheless, we claim that the program of Quine has succeeded, at least as well as the program of axiomatic set theory in the style of Zermelo. But nobody noticed. The proof of its success came quite a while ago, in 1969, when R. B. Jensen proved in [19] the consistency of what seemed a slight modification of *NF*; he restricted extensionality to objects with elements, so as to allow urelements, obtaining the system which he called *NFU*. In Zermelo-style set theory, to have urelements or not is a technical decision, decided for technical reasons in favor of strong extensionality; in Quine-style set theory, the effects are profound. *NFU* is not only consistent (its consistency is a theorem of Peano arithmetic), but consistent with the Axiom of Choice as well. If the theory is so weak, why is it interesting? Jensen showed in addition that *NFU* with the Axioms of Infinity and Choice is consistent; this is a theory exactly as strong as the theory of types with the Axiom of Infinity or bounded Zermelo set theory (Zermelo set theory with only the Δ_0 instances of comprehension, those in which each quantifier is restricted to a set). *NFU* + Infinity + Choice is as easy to use as either of these theories. Jensen went on to show that for each ordinal

α there is a model of *NFU* + Choice which contains a well-ordering of type α; this implies, for instance, that there are models of *NFU* with Choice in which all natural numbers are standard. This implies that *NFU* admits extensions as strong as any extension of *ZFC* in which we have confidence (roughly speaking).

In 1953, interest in *NF* as a foundation for mathematics was strong enough that J. B. Rosser could choose this system as the medium for the development of the foundations of mathematics in his *Logic for Mathematicians* ([27]). While the system of Rosser's book (*NF* + "the axiom of counting" (discussed below) + denumerable choice) has not been shown to be inconsistent, the result of Specker that the full Axiom of Choice fails caused mathematicians to lose interest in applying this system. The system *NFU* + Counting + (full) Choice is consistent relative to *ZFC*, and the development in *Logic for Mathematicians* can be duplicated there, except that a different way of introducing the ordered pair is needed (for a reader of the book, the easiest approach is to treat Rosser's pair as a primitive notion). By the time that this was shown by Jensen in 1969, there was no interest in doing this; in fact, the consistency results for *NFU* seem not to be well-known.

The program of Quine has spawned other set theories now known to be consistent, notably the theory NF_3 (extensionality + stratified comprehension using only three types) shown to be consistent by Grishin in [10], also in 1969, and the theory *NFI* (an extension of predicative *NF*: extensionality + stratified comprehension with no type used which is higher than that which would be assigned to the set being defined) shown to be consistent by Marcel Crabbé in [4] in 1982. The model theory developed by Specker for *NF* (in [31]) has nontrivial applications in the model theory of *NFU*, NF_3, and *NFI*. NF_3, *NFI*, and those of their extensions which are known to be consistent have independent interest, although they are not, as suitable extensions of *NFU* are, well-suited for use as "all-purpose set theories" in the same way as *ZFC*.

ZFC works reasonably well as the "all-purpose set theory" at the foundations of mathematics (the need for proper classes is often felt). Why should we bother considering an alternative theory? First, extensions of *NFU* have the property of admitting all the constructions of naive set theory except those which involve induction on the membership relation; for example, objects like the universal set and the Frege natural number 3, the set of all sets with three elements, exist. Most natural categories are sets in *NFU* (although, as Colin McLarty has pointed out in his note [21], the properties of these "big" categories are not quite what one would expect). The von Neumann construction of the ordinals involves induction on the membership relation and so fails; but the original definition of ordinal numbers as equivalence classes of well-orderings under similarity succeeds in *NFU*. The usual inductive construction of the universe in ordinal-indexed stages also depends on induction on membership, but it turns out that this, too, can be modelled in *NFU* using equivalence classes of well-founded extensional relations. Moreover, it can be arranged for these constructions to succeed on a limited domain (not a set) which will look like the universe of a Zermelo-style set theory, without adding consistency strength. There is a "natural" extension of *NFU* in which *ZFC* can be interpreted in this way.

ZFC has an intuitive motivation; the construction of the universe in iterative stages is a convincing metaphor helping us to regard the theory and certain of its extensions as reliable. For a recent discussion of this, see [20]. *NF* was originally proposed as a "syntactical trick," and so seems less reliable; we will present an intuitive argument for the "reasonableness" of the stratification criterion for comprehension. The "intuitive" motivation we give for Quine-style theories seems to shed more light on the reasons for the paradoxes of naive set theory than the "limitation of size" motivation for *ZFC*, and so we are able to excise strictly less of naive set theory than we do with the Zermelo approach. The consistency results for extensions of *NFU* show that we are not being too liberal.

It is our contention that *NF* itself is not interesting as a foundation for mathematics. *NFU* (with Choice and strong axioms of infinity) is the "useful" theory of this type. But it remains an intriguing mathematical question why this is so. Why does the strong axiom of extensionality, harmless in *ZFC*, have such profound effects when adjoined to *NFU*? We believe that we know how to answer this question; we can exhibit a version of the theory of types and an extension of bounded Zermelo set theory which are equiconsistent with *NF* and should help the reader to see why *NF* is so strange. The consistency question for *NF* remains, but Crabbé's fragment *NFI*, which has much the same strange character, is known to be consistent. The obstruction to extending this result has to do with the impredicativity of *NF*. We believe that Specker has already shown that *NF* is not "useful" by showing that Choice fails there.

Another program for the foundations of mathematics interacts fruitfully with the Quine program. Curry's attempt to found mathematics on type-free combinatory logic (see [6]) can be put on a firm footing using restrictions on function abstraction analogous to the stratification restrictions on set comprehension in *NF*. We have constructed systems of combinatory logic equivalent to *NF* and *NFU* + Infinity (in [13]) and to *NFI* (in [15]) which may have practical applications in computer science. Our argument for the "reasonableness" of stratified comprehension, translated into the context of combinatory logic or λ-calculus, turns out to involve concepts analogous to well-known ideas in computer science. Intuitionistic *NF* is being investigated (see [7]) but little is known about it as yet.

3 The Historical Motivation of *NF*

NF was introduced by W. V. Quine in 1935 in his paper [24]. His original motivation for defining *NF* was strictly syntactical, arising from the phenomenon of "typical ambiguity" in Russell's Theory of Types (introduced in [28], although in a more complicated form than that presented here). So we will begin by discussing the Theory of Types. As we will see below, it is possible to motivate *NF* (or *NFU*, at least) semantically rather than syntactically, but the theory has gone through most of its history innocent of semantic motivation.

We present the Theory of Types (TT) in a modern formulation. TT is a first-order many-sorted theory with equality and membership. The sorts are indexed by the non-negative integers $0, 1, 2\ldots$; they are called "types". Objects of type 0 are called "individuals"; objects of type 1 are to be understood as sets of individuals; objects of type 2 are to be understood as sets of type 1 objects (sets of sets of individuals); in general, a type $n+1$ object is a set of type n objects. In keeping with the indicated interpretation, a formula $x \in y$ is well-formed if and only if the type of y is the successor of the type of x; a sentence $x = y$ is well-formed if and only if x and y have the same type. It should be noted that the the types are not "disjoint" in any sense expressible in the theory; it is impossible to formulate the assertion that objects of different types are either equal or unequal. It is also useful to note that type 0 objects should not be thought of as urelements; to ascribe elements to a type 0 object is meaningless rather than false.

The axiom schemes of TT are extensionality (objects of a positive type $n+1$ are equal if they have the same type n members) and comprehension (for each formula ϕ and variable x of type n there is an object $\{x \mid \phi\}$ of type $n+1$ such that for all x of type n, $x \in \{x \mid \phi\}$ iff ϕ). It is usual to add the Axiom of Infinity and perhaps the Axiom of Choice, but we do not regard these as part of the basic theory.

We illustrate the phenomenon of "typical ambiguity" which motivated the definition of *NF* for Quine. Suppose we want to define the number 3. In TT, we can use Frege's definition: 3 is the set of all sets with three elements. This is not a circular definition, since we can express the property of having three elements in first-order logic with equality and membership. But observe that we need a different number 3 to count objects in different types; the number 3 which is the collection of all sets of three type n objects lives itself in type $n+2$. The arithmetic of all these numbers 3 is exactly parallel; it is philosophically unappealing to have so many copies of what "should" be a single object.

This phenomenon is perfectly general. For any formula ϕ in the language of TT, we define ϕ^+ as the formula which results if we raise each type index in ϕ by one. Now observe that if ϕ is an axiom, so is ϕ^+. It follows immediately that if ϕ is a theorem, so is ϕ^+. Any object $\{x \mid \phi\}$ that we define has a sequence of analogues $\{x^+ \mid \phi^+\}$, $\{x^{++} \mid \phi^{++}\}$, and so on, one in each higher type.

Quine's suggestion was that the types are redundant; we should understand all the various versions of the number 3 as being the same object, and, in general, we should regard the sequence of objects $\{x \mid \phi\}$, $\{x^+ \mid \phi^+\}$, $\{x^{++} \mid \phi^{++}\}, \ldots$ as being one untyped object $\{x \mid \phi\}$. The danger of paradox is apparently avoided by only allowing those formulas ϕ to define sets which would make sense in the theory of types with a suitable assignment of types to their variables; we move to a type-free theory, but we only allow those instances of comprehension which are type-free analogues of instances of comprehension in TT. For example, the Russell class $\{x \mid x \notin x\}$ is avoided because there is no analogous instance of comprehension in TT, but the universe $V = \{x \mid x = x\}$ does exist, as does Frege's number 3 (now found in only one version).

The notion of an instance of untyped comprehension being analogous to an instance of comprehension of TT needs to be formulated precisely. A formula ϕ in the language of one-sorted first-order logic with equality and membership is said to be "stratified" if it is possible to assign a non-negative integer to each variable occurring in ϕ in such a way that each variable always appears with the same associated integer, each subformula $x \in y$ has successive integers assigned to x and y and each subformula $x = y$ has the same integer assigned to x and y. It should be clear that ϕ is stratified exactly when a suitable assignment of types to variables would convert it to a well-formed formula of TT.

With the notion of stratification defined, we can define Quine's set theory *NF*. *NF* is the one-sorted first-order theory with equality and membership whose axioms are extensionality (objects with the same elements are equal) and the scheme of stratified comprehension ($\{x \mid \phi\}$ exists, for each stratified ϕ and variable x). We do not add Infinity or Choice; as we will see below, Infinity is a difficult *theorem* of this system, while Choice is false! It is worthwhile to note an error in Quine's original paper [24]: Quine claimed that *NF* proved Infinity because of the existence of each of the objects $\{\}, \{\{\}\}, \{\{\{\}\}\} \ldots$; but the collection of iterated singletons of the empty set does not have a stratified definition and cannot be proven to exist. While Infinity is a theorem of *NF* itself for different reasons, the related theory *NFU* (*NF* with urelements) proves the existence of each of the above objects, but has models in which the cardinality of V is a (nonstandard) natural number.

Specker, in [31], put Quine's intuition about the way in which *NF* is obtained from TT by collapsing the type structure on a firm footing. He showed that the following are equivalent:

a. *NF* is consistent (i.e., there is a model of *NF*).

b. TT + the axiom scheme "$\phi \iff \phi^+$" for all formulas ϕ (the axiom scheme of *typical ambiguity*) is consistent.

c. There is a model of TT with a "type-shifting automorphism" (a bijective map which sends type n onto type $n+1$ for each n and respects membership).

These model-theoretic results show that Quine's intuitive motivation for *NF* can be precisely formalized. As yet, no one has been able to show the consistency of TT + Ambiguity or build a model of TT with a type-shifting automorphism; the question of the consistency of *NF* remains open. But these techniques also apply to fragments of *NF* which are known to be consistent, relating them to ambiguous versions of related fragments of TT. We will present an argument below with the same flavour as Specker's proofs of the results cited above.

It is interesting to observe that a weak version of the usual set theory can be derived by collapsing the type distinctions in Russell's theory in a different way (although we do not claim that this is the historical motivation for set theory in the style of Zermelo). Suppose that we are given an injection of type 0 (individuals) into type 1 (sets of individuals), witnessed internally to TT by a bijection on type 1 with domain the set of singletons of objects of type 0; the identity map on these singletons is the simplest example. We can view this map as an *identification* of type 0 with part of type 1. We can then identify each type 1 set with the type 2 set obtained by replacing each of its type 0 elements with the type 1 object with which it is identified, and continue inductively, identifying each type $n+1$ set with the type $n+2$ set which is obtained by replacing each of its type n elements with the type $n+1$ object identified with it. The identification of type n with type $n+1$ is in each case witnessed by a map definable in TT, which takes the singleton of each type n object to the type $n+1$ object with which it is identified. We can then construct a direct limit, collapsing all the types into one domain. This domain will satisfy the axioms of Zermelo set theory without foundation (although all failures of foundation are accounted for by the bad behaviour of a set of objects, the original type 0) and with the restriction on the axiom scheme of Comprehension that all quantifiers must be restricted to a set (Δ_0 comprehension). Infinity and Choice will be satisfied if they held in the original TT. The reason for this restriction is that any attempt to define a set in this theory must be carried out below some type of TT, itself interpreted as a set which can be taken to bound all quantifiers in the set definition. The axiom of foundation can be recovered; instead of taking the base of the construction to be the original type 0, let it be the set of natural numbers in type 2, and interpret higher types accordingly; use the map which sends n to the set of numbers less than n as the embedding of "type 0" in "type 1," so that the type 0 objects are interpreted as von Neumann ordinals and Foundation will hold. This construction witnesses the fact that TT + Infinity is equiconsistent with Zermelo set theory with Δ_0 comprehension, which we will call "bounded Zermelo set theory". The "disjoint" types of Russell's theory can be associated with a sequence of consecutive cumulative types in a Zermelo-style theory, but nothing in the Russell theory corresponds to what happens at cumulative types indexed by limit ordinals in a Zermelo-style theory.

4 A Survey of Jensen's Theory *NFU*

In 1969, in [19], R. B. Jensen described the theory *NFU* ("New Foundations with urelements") whose difference from *NF* is that it allows "atoms" or "urelements" (objects with no elements which are allowed to be distinct from one another). In [24], Quine had suggested that the choice of strong extensionality for *NF* was merely a technical decision; he suggested that extensionality could be extended to atoms by redefining membership so that they could be taken to be their own sole elements. This turned out to be incorrect, as the proposal overlooked the effects of stratification. The difference between *NF* and *NFU* is profound.

Jensen was able to show that *NFU* is consistent (this is a theorem of Peano arithmetic!) and that it can be extended with the axioms of Infinity and Choice (unlike *NF*, it neither proves the former nor disproves the latter). Moreover, he was able to construct ω-models of *NFU* (models in which all natural numbers are standard) and to do this in a way which could be generalized; Jensen showed that for each ordinal α, there is a model of *NFU* + Infinity + Choice which contains a well-ordering of type α (he showed this in *ZFC*). Jensen's result shows that *NFU* has consistent extensions of essentially the same strength as any extension of the usual set theory in which we have confidence.

In his remarks accompanying [19], Quine suggested a convenient form for the axiomatization of *NFU*; one wants to distinguish the empty set from the other urelements which are non-sets, so one provides a primitive sethood predicate. The axiom of extensionality is then modified to assert that any two *sets* with the same elements are equal, while an axiom of atoms asserts that non-sets have no elements. The axiom of stratified comprehension asserts that for each stratified formula ϕ, there is a *set* $\{x \mid \phi\}$ ("formula" is correct here in place of "sentence" because parameters can appear in set definitions). Quine suggests also that it was an error on his part to assume that it was harmless to suppose in the unfamiliar context of a theory with stratified comprehension that all objects are sets.

We exhibit a version of Jensen's proof of Con(*NFU*). We will prove the consistency of TTU (the theory of types with urelements allowed in each type) with the scheme of typical ambiguity. Adapting Specker's results on the model theory of *NF* in [31], this is equivalent to the consistency of *NFU*. Below, we shall construct a model of TTU with a type-shifting automorphism, and indicate how it can be converted into a model of *NFU*.

Let V_α be stage α in the iterative construction of the universe of the usual set theory, for each ordinal α. For each pair of ordinals $\alpha < \beta$, define "$x \in_{\alpha,\beta} y$" as "$x \in y$ and $y \in V_{\alpha+1}$," for each x, y in V_β. We can then use any ascending sequence of ordinals α_i to determine a model of TTU: let type i be $V_{\alpha_i} \times \{i\}$, and define "$(x, i) \in (y, i+1)$" as "$x \in_{\alpha_i,\alpha_{i+1}} y$". Observe that elements of $V_{\alpha_{i+1}} - V_{\alpha_i+1}$ are interpreted as urelements under the "membership" relation $\in_{\alpha_i,\alpha_{i+1}}$.

Let Σ be a finite set of formulae of the language of TTU. Let $n-1$ be the highest type index which appears in Σ. Fix a limit ordinal λ. One can associate with Σ a partition of the set of n-element sets of ordinals less than λ, determined by the truth-values of the formulas in Σ when the sequence of types in an n-element set is used to interpret types 0 through n of TTU as indicated in the paragraph above. Use the Ramsey theorem to choose an infinite homogeneous set H for this partitition; the interpretation of TTU using any ascending sequence of ordinals in H will satisfy the scheme of typical ambiguity $\phi \iff \phi^+$ for all formulas ϕ in Σ. By compactness, we conclude that the scheme of typical ambiguity is consistent with TTU, from which it follows by the model theory results of Specker that *NFU* is consistent. It is easy to see that Infinity and Choice will hold in the interpretation of TTU determined in this way if we use infinite ordinals as types and if choice holds in the underlying version of the usual set theory; it follows easily that Infinity and Choice are consistent with NFU. Jensen constructed ω- and α-models of NFU using a more sophisticated argument of this sort involving the Erdös-Rado theorem.

We show how to construct a model of *NFU* directly. Work in a nonstandard model of set theory which has an external automorphism j and an ordinal α such that $j(\alpha) > \alpha$. This gives us a model of TTU with a shifting automorphism, determined as indicated above by the sequence of ordinals $\alpha, j(\alpha), j^2(\alpha), \ldots$. It is not a problem that this sequence is external to the model, since constructions in TTU never involve more than finitely many types. We now build a model of *NFU* directly as follows: let V_α be the collection of objects of the model, and define the membership $\in'$ of the model as follows: $x \in' y$ if and only if $x \in j(y)$ and $j(y) \in V_{\alpha+1}$ (or, equivalently, and referring only to objects in the model, $j^{-1}(x) \in y$ and $y \in V_{j^{-1}(\alpha)+1}$). It is easy to show that the axioms of Infinity and Choice hold in such a model if they hold in the underlying model of set theory (if α is nonstandard finite, we get a model of NFU + "the universe is finite"). The elements of $V_\alpha - V_{j^{-1}(\alpha)+1}$ are interpreted as urelements.

We prove that the model described above is actually a model of NFU. The argument we give has the same flavor as Specker's proofs of his model-theoretic results cited above. Let ϕ be a stratified formula in the language of first-order logic with equality and membership. Translate it into a formula ϕ' in the language of equality, membership, and the automorphism j by translating each occurrence of membership "$x \in y$" as "$x \in j(y)$ and $j(y) \in V_{\alpha+1}$"; also, restrict all quantifiers to V_α. We have moved from the language of *NFU* to the language of the nonstandard model of the usual set theory with the automorphism j. The obstruction to finding the set $\{x \mid \phi'\}$ in the latter system is the fact that j is external; we need to show that j can be eliminated from ϕ'. Observe that assertions "$x = y$" and assertions "$x \in y$" are precisely equivalent to "$j^i(x) = j^i(y)$" and "$j^i(x) \in j^i(y)$" respectively. Apply the isomorphism j to both sides of each atomic formula in ϕ' in such a way that each variable z assigned type i in the stratification of ϕ appears only in the context $j^i(z)$ (itself not in the scope of further applications of j or its inverse). It should be clear that this is possible, since the translation of $\in$ and the fact that $=$ is left unperturbed by our original translation means that the conditions on applications of j to variables in atomic subformulae of ϕ' are exactly that the exponents on both sides of an $=$ are the same, while the exponents on the left and right sides of an $\in$ are successive; these conditions are preserved by applications of j to both sides of atomic subformulae and correlate exactly with the conditions for type assignments in stratified formulae. Now we can replace $j^i(z)$, for each variable z assigned type i, with a variable z restricted to $j^i(V_\alpha)$ instead of V_α; this has the effect that j no longer occurs except in constants, which means that $\{x \mid \phi'\}$ actually exists.

The set $\{x \mid \phi\}$ in the model of *NFU* is the set $j^{-1}(\{x \mid \phi'\})$ in the nonstandard model of set theory. Observe that if ϕ had not been stratified, we would not have been able to eliminate essential occurrences of j in ϕ', and we would not have been able to conclude that the latter formula had an extension.

We will now consider how the foundations of mathematics could be laid in an extension of *NFU*. We will present a finite axiomatization of *NFU* + Infinity + Choice, modelled on von Neumann's axioms for constructing proper classes over the usual set theory combined with ideas of Grishin (see [10]) on *NF*. This axiomatization should help to dispel the idea that *NFU* is based merely upon a syntactical trick; like the usual axiomatization of *ZFC*, it shows how to build sets using basic sets and constructions on sets which seem intuitively legitimate, and does not mention the syntactical notion of "stratification". The language of *NFU* is augmented for our present purposes with a primitive ordered pair; we will discuss the reasons for this below. Although we will express the axioms using notation like "(x, y)," the official extension of our language adds relations π_1 and π_2, the relations in which a pair stands to its first and second projections, respectively. For purposes of stratification, the variables in a formula $x\pi_i y$ should be assigned the same type.

Note: The usual notation for sets and relations defined as extensions of formulas is used freely in the statement of the axioms; it can, of course, be eliminated.

Axiom 1. Sets with the same elements are equal.

Axiom 2. Objects which are not sets have no elements.

These axioms reiterate the treatment of extensionality in *NFU* outlined above. It will be a theorem of this system that there are non-sets; this theorem will be a consequence of the Axiom of Choice.

Axiom 3. The universe $V = \{x \mid x = x\}$ exists.

Axiom 4. For each set A, the set $A^c = \{x \mid x \notin A\}$ exists.

Axiom 5. For each pair of sets A and B, the set $A \cup B = \{x \mid x \in A \text{ or } x \in B\}$ exists.

These axioms assert that the universe of sets forms a Boolean algebra with the usual operations of complement (provided by Axiom 4), union (provided by Axiom 5), and intersection (definable in terms of the other two as usual). Axiom 3 is actually redundant, since axioms below will provide for the existence of specific sets, but it is useful to give it here for the sake of orderly development. We can now prove the existence of $\{\}$, the unique *set* with no elements.

Axiom 6. For each x, the set $\{x\} = \{y \mid y = x\}$ exists.

Axiom 7. If $(x, z) = (y, w)$, then $x = y$ and $z = w$.

Axioms 6 and 7 give us the ability to construct finite ordered and unordered structures. Application of Axiom 6 and the axiom of Boolean union above enables us to construct finite sets; axiom 7 allows us to construct n-tuples, which we define inductively so that $(x_1, \ldots, x_n) = (x_1, (x_2, \ldots, x_n))$.

Axiom 8. For each pair of sets A and B, the set $A \times B = \{(x, y) \mid x \in A$ and $y \in B\}$ exists.

We define a *relation* as a set of ordered pairs and read xRy as $(x, y) \in R$ as usual.

Axiom 9. For each relation R, the set $\mathrm{dom}(R) = \{x \mid \text{for some } y, xRy\}$ exists.

Axiom 10. For each relation R, the relation $R^{-1} = \{(x, y) \mid yRx\}$ exists.

Axiom 11. For each pair of relations R and S, the relation $R \mid S = \{(x, y) \mid$ for some z, xRz and $zRy\}$ exists.

Axiom 12. The relation $=$, $\{(x, x) \mid x \in V\}$, the relation π_1, $\{((x, y), x) \mid x, y \in V\}$, and the relation π_2, $\{((x, y), y) \mid x, y \in V\}$, exist.

Axioms 8 through 12 provide us with the machinery for a theory of relations. From Axioms 1 through 12, it is possible to prove the following:

Meta-Theorem 1: Let ϕ be a formula of first-order logic with equality such that for each atomic n-place predicate R occurring in ϕ the set $\{(x_1, \ldots, x_n) \mid Rx_1 \ldots x_n\}$ exists. Then $\{x \mid \phi\}$ exists.

Indication of Proof: Using Axioms 1 through 12, it is possible to define the operations of Tarski's cylindrical algebra (see [12]) on n-place relations for each n; this can be used to define the set of n-tuples which satisfy any given formula of first-order logic, as long as each atomic predicate appearing in the given formula defines a set. It is possible to define projections of sets of n-tuples on any coordinate, and this gives us the ability to define any set $\{x \mid \phi\}$ if the atomic relations used in ϕ are available as sets.

From the Meta-Theorem, we can prove the following discouraging theorem:

Theorem: The membership relation $\{(x, y) \mid x \in y\}$ does not exist.

Proof: If this relation existed, the Russell class $\{x \mid x \notin x\}$ would exist by Meta-Theorem 1. More concretely, if the set above existed (call it E for the nonce), the set $R = \mathrm{dom}(E^c \cap =)$, the Russell class, would exist.

The axiomatization of *NFU* + Infinity is completed as follows:

Axiom 13. The relation $\subseteq\, = \{(x, y) \mid x \subseteq y\}$ exists.

Axiom 14. For each set A, the set $\bigcup A = \{x \mid \text{for some } y, x \in y \text{ and } y \in A\}$ exists.

Axiom 15. For each relation R, the relation SI$\{R\}$ (SI for "singleton image") $= \{(\{x\}, \{y\}) \mid xRy\}$ exists.

The set $\mathcal{P}_1\{A\} = \{\{x\} \mid x \in A\}$ (the "singleton image" of the *set* A) can be defined as dom(SI$\{A \times \{A\}$). Since we cannot have the membership relation, we settle for the inclusion relation. Along with the machinery of set unions and "singleton images" of sets and relations, this enables us to prove:

Meta-Theorem 2: If ϕ is a stratified formula in the language of *NFU* as extended here, then $\{x \mid \phi\}$ exists.

Indication of Proof: We use a technique due to Grishin to eliminate the "non-existent" membership relation from ϕ. Replace "$x \in y$" with "$\{x\} \subseteq y$". We demonstrate the existence, not of the set $\{x \mid \phi\}$ directly, but of a suitable iterated singleton image of this set, by applying the singleton operation to both sides of each atomic formula of ϕ until a variable y assigned type i in the stratification of ϕ appears only in its $(N - i)$-fold singleton, and not in its $(N - i + 1)$-fold singleton, for N a suitably large natural number; this is possible because ϕ is stratified and the applications of the singleton operation in the translation of $\in$ on the left of an inclusion symbol exceed those on the right by one. The $(N - i)$-fold singleton of y can then be replaced by a variable y restricted to the set of all $(N - i)$-fold singletons, $\mathcal{P}_1^{N-i}\{V\}$, and the atomic sentences still determine sets because inclusion, equality, and projection relations are sets and iterated singleton images of these relations are sets. Meta-Theorem 1 is then applied to establish the existence of $\mathcal{P}_1^j\{\{x \mid \phi\}\}$ for some j (the number of applications of the singleton operation with which x appears), and the axiom of set union is applied j times to yield the existence of $\{x \mid \phi\}$. Unstratified formulas ϕ cannot be relied upon to have extensions because the relation between objects and their singletons is not provided. Note the similarity of the proof to the verification of our model of *NFU* above.

It will turn out that we do not have just *NFU* here, but *NFU* + Infinity. The reason is that we have elected to introduce a primitive ordered pair with the same type as its projections, instead of using the definable Kuratowski pair $< x, y > = \{\{x\}, \{x, y\}\}$, which is two types higher than its projections. There is actually a definable pair in *NFU* which is *one* type higher than its projections. The original finite axiomatization of *NF* due to Hailperin used the Kuratowski pair, and could be adapted to *NFU* without Infinity, but it is much harder to understand than the axiomatization given here. Our axiomatization can also be adapted to *NFU* without Infinity, but it makes it harder to work with, and we will not give the details here. It should be noted that it is not actually possible to prove the existence of a type-level ordered pair in NFU + Infinity, so the theory here is inessentially stronger than plain NFU + Infinity; "inessentially" stronger because it is easy to interpret NFU + "there is a type-level ordered pair" in NFU + Infinity; it should also be noted that one can prove the existence of a type-level ordered pair in NFU + Infinity + Choice.

Now that we have Meta-Theorem 2, we will use stratified comprehension freely. We will do this very informally; the best way to verify to oneself that a construction is stratified is to verify that it can be carried out in type theory.

Arithmetic in NFU follows Frege. It is most natural to start with some general cardinal arithmetic. The notion of equivalence of A and B is defined as usual as obtaining exactly when there is a bijection between A and B. We can then define $|A|$, the cardinal number of the set A, as the set of all sets equivalent to A. We define 0 as $|\{\}|$ (note that the empty set is the sole element of 0; the nonsets are not included), and 1 as $|0|$ (any other singleton would do!). We then define addition and multiplication: $|A|+|B| = |(A \times \{0\}) \cup (B \times \{1\})|$ and $|A||B| = |A \times B|$.

We can now define the natural numbers as the finite cardinals: we call a set of cardinal numbers *inductive* if it contains 0 and is closed under addition of 1. The set $\mathcal{N}$ of natural numbers can then be defined as the intersection of all inductive sets.

The axioms of Peano arithmetic are easily verified (closure of $\mathcal{N}$ under addition and multiplication as defined needs proof, of course); we omit this, but indicate how the Axiom of Infinity is proven. If the universe were finite, V would be an element of some natural number n, and $n+1$ would be empty (we can prove by induction that no set belonging to a natural number has a proper subset belonging to the same natural number). We can prove for any natural number m which is nonempty that $m+1$ is nonempty; in fact, the definition of addition allows us to exhibit an element of $\kappa+1$ given an element of κ for any cardinal number κ! By induction (since 0 is certainly nonempty), all natural numbers are nonempty and V can belong to no natural number. The presence of the type-level pair is what makes this possible; if the pair were not at the same type as its projections, $A \times \{0\}$ would not necessarily have the same cardinal as A (the definition of the obvious bijection would be unstratified), and the definition of addition as we have given it would fail.

Defining the exponential function in the natural way (for general cardinals first, then for natural numbers, as above), presents some problems. We would like to define $|B|^{|A|}$ as the cardinality of the set B^A of functions from A to B. This definition succeeds, in the sense that it does define an operation on cardinals. However, if this operation is carried out in TT, the cardinal number $|B|^{|A|}$ defined in this way is one type higher than $|A|$ or $|B|$. As a result, we cannot define the function which takes $(|B|, |A|)$ to $|B|^{|A|}$; its definition is not stratified. We cannot define the map which takes $|A|$ to $2^{|A|}$, either. Another problem is that the number $|A|^2$ defined in this way is not necessarily equal to $|A| \times |A|$.

There is a "trick" which enables us to cope with this problem. For each cardinal $|A|$, define $\mathrm{T}\{A\}$ as $|\mathcal{P}_1\{A\}|$, the cardinality of the singleton image of A. T is not a function; $\mathrm{T}\{\kappa\}$ is one type higher than κ. It is easy to show that this is a well-defined operation on cardinals, and that its inverse is a well-defined partial operation on cardinals. We define $|B|^{|A|}$ as $\mathrm{T}^{-1}\{|B^A|\}$; the exponential $|B|^{|A|}$ is now of the same type as $|A|$ and $|B|$, but it is not always defined, as we will see below. The T operation has very nice arithmetic properties: we can prove that $\mathrm{T}\{\kappa\} + \mathrm{T}\{\lambda\} = \mathrm{T}\{\kappa + \lambda\}$, that $\mathrm{T}\{\kappa\}\mathrm{T}\{\lambda\} = \mathrm{T}\{\kappa\lambda\}$, and that $\mathrm{T}\{\kappa\}^{T\{\lambda\}} = \mathrm{T}\{\kappa^\lambda\}$. We will see below that it is nonetheless nontrivial.

In the special context of the natural numbers, we can prove that n^m is defined and a natural number for all n and m by an inductive argument. It is easy to see that $\mathrm{T}\{0\} = 0$, and it is also easy to prove that $\mathrm{T}\{n+1\} = n+1$ if $\mathrm{T}\{n\} = n$; but it is not possible to prove by induction that $\mathrm{T}\{n\} = n$ for all natural numbers n; $\mathcal{N}$ is defined as the intersection of all inductive *sets*, so induction succeeds for any formula ϕ which is inductive and has an extension; "$\mathrm{T}\{n\} = n$" is an unstratified formula and need not have an extension. It is known that the assertion that $\mathrm{T}\{n\} = n$ for all n cannot be proven from the axioms above, but it is also known that it is consistent with the axioms above. We introduce it, therefore, as

Axiom 17. For each natural number n, $\mathrm{T}\{n\} = n$.

This axiom, due to Rosser (in [27], where it was proposed in the context of *NF*), is called the Axiom of Counting. It is easy to show that the set $\{1, \ldots, n\}$ of the first n positive natural numbers belongs to $\mathrm{T}\{n\}$; thus Axiom 17 justifies the intuitive idea of counting using the positive integers (Rosser's original form for the axiom was "$\{1, \ldots, n\} \in n$," which is equivalent). Our model construction above will yield a model of *NFU* + Counting exactly when the automorphism j moves no natural number. Under the Axiom of Counting, the two definitions for exponentiation proposed in turn above coincide for natural numbers.

Note that, although induction for unstratified conditions is not a consequence of our axioms, this does not indicate any weakness in the arithmetic of our theory; all formulae of the interpretation of Peano arithmetic (or, indeed, of nth order arithmetic for any concrete n) are stratified. Jensen showed in [19] that *NFU* has ω-models, which demonstrates that it is consistent with full mathematical induction; we will introduce a strong axiom which will imply full mathematical induction (and much more) below.

We now consider general cardinal arithmetic. The definition of order relations for cardinal numbers is the usual one, and the Schröder-Bernstein theorem can be proven for infinite cardinals much as in the usual set theory.

We have the theorem of Cantor, $2^{|A|} > |A|$, but it does not mean quite the same thing as it does in the usual set theory. The usual form $|A| < |\mathcal{P}\{A\}|$ of this theorem is obviously false here, because $|V|$ is certainly not less than

$|\mathcal{P}\{V\}|$. The power set of the universe, $\mathcal{P}\{V\}$, is the set of all subsets of the universe, or the set of all *sets*. We will see below that it follows from the Axiom of Choice that $|\mathcal{P}\{V\}|$ is considerably smaller than $|V|$; there are many urelements! However, $2^{|A|}$ has not been defined as $|\mathcal{P}\{A\}|$, but as $\mathrm{T}^{-1}\{|\{\mathcal{P}\{A\}|\}$.

Cantor's original proof that $|A| < |\mathcal{P}\{A\}|$ went as follows: suppose that f is a bijection from A onto $\mathcal{P}\{A\}$ ("clearly" $|A| \leq |\mathcal{P}\{A\}|$, so the only possibility which needs to be eliminated is exact equality). Define the set R as $\{x \in A \mid x \notin f(x)\}$; now consider whether $f^{-1}(R) \in R$. This fails in *NFU* because the defining condition for R is not stratified. The analogous argument in *TT* or *NFU* shows that $|\mathcal{P}_1\{A\}| < |\mathcal{P}\{A\}|$. The corresponding nonstrict inequality clearly holds (it is not necessarily the case that $|A| \leq |\mathcal{P}\{A\}|$ in *NFU*, by the way). Let f be a bijective map from $\mathcal{P}_1\{A\}$ onto $\mathcal{P}\{A\}$; consider the set R defined as $\{x \in A \mid x \notin f(\{x\})\}$, and ask whether $f^{-1}(R) \subseteq R$; a contradiction follows from the existence of such an f as in the original argument. Now it is the case that $2^{|\mathcal{P}_1\{A\}|} = |\mathcal{P}\{A\}|$, so we have proven a special case of the desired inequality on cardinals (the case in which the exponent is the cardinality of a set of singletons); commutativity of T^{-1} with arithmetic operations turns out to be sufficient to prove the general case from this.

In the special case $A = V$, this establishes that $|\mathcal{P}_1\{V\}| < |\mathcal{P}\{V\}| \leq |V|$. This means that the unrestricted singleton map (ι such that $\iota(x) = \{x\}$ for all x) cannot exist in *NFU*; this should not be surprising, considering its unstratified definition and its role in the proof of Meta-Theorem 2 above. For any set A such that $|A| = |\mathcal{P}_1\{A\}|$, the Cantor theorem holds in its original form; such sets are called "Cantorian" sets and their cardinals are called Cantorian cardinals; the Cantorian cardinals are the fixed points of the T operation. Observe that Cantor's theorem implies that $|V|$ is not Cantorian, and so that the T operation on cardinals is nontrivial, and that the Axiom of Counting asserts that all finite cardinals are Cantorian. A set A such that $(\iota|A)$ (the singleton map restricted to A) exists is called a "strongly Cantorian" set, and the cardinal of such a set is called a strongly Cantorian cardinal. All strongly Cantorian sets are Cantorian; the converse is a strong axiom of infinity. $|\mathcal{N}| = \aleph_0$ can be proven to be Cantorian in *NFU* + Infinity; the assertion that $\aleph_0$ is strongly Cantorian is equivalent to the Axiom of Counting. The conditions "Cantorian" and "strongly Cantorian" are defined by unstratified formulae and do not determine sets. The hereditarily strongly Cantorian sets can be thought of as analogous to the "small" sets of the usual set theory; we will return to this.

We can subvert the stratification restrictions on comprehension with respect to variables restricted to a strongly Cantorian set. The trick is to use $(\iota|A)$ (the singleton map restricted to A) to raise or lower the type of a variable x restricted to A. We can raise the type of x by replacing it with "the element of $(\iota|A)(x)$," and lower the type of x by replacing it with $(\iota|A)^{-1}(\{x\})$. For example, the set $\{x \in A|x \notin x\}$ can be defined as $\{x \in A|(\iota|A)(x) \not\subseteq x\}$, if A is strongly Cantorian. Note that the Axiom of Counting allows us to apply this; we do not need to consider the types of variables restricted to $\mathcal{N}$. The map $(\iota|\mathcal{N})$ is defined by induction: $(\iota|\mathcal{N})(0) = \{0\}$ and $(\iota|\mathcal{N})(m+1) = \{n+1\}$ if $(\iota|\mathcal{N})(m) = \{n\}$; this is a stratified definition, which can be shown in *NFU* + Infinity to be a bijection between $\mathcal{N}$ and $\mathcal{P}_1\{\mathcal{N}\}$; the Axiom of Counting is needed to show that $f(n) = \{n\}$ for each n in $\mathcal{N}$.

In the theory of ordinals, as in the theory of cardinals, our definitions can follow those of naive set theory. An ordinal number is an equivalence class of well-orderings under similarity. The natural order on the ordinals ($\alpha < \beta$ if and only if an element of α is similar to an initial segment of an element of β) is easily shown to be a well-ordering. Transfinite induction works for stratified conditions. Since the natural order on the ordinals is a well-ordering, it belongs to an ordinal Ω. In naive set theory, we would now have the Burali-Forti paradox: the order type of the ordinals below Ω in the natural order could be proven to be Ω itself, which would imply that the natural order on the ordinals was similar to a proper initial segment of itself, which is impossible. The point at which this proof fails in *NFU* is the assertion that the order type of the ordinals below Ω is Ω. The assertion "the order type of the ordinals below α is α for each ordinal α" can be proven by transfinite induction in the usual set theory, but it cannot be proven by induction in *NFU*, because the induction is on an unstratified condition, which may not define a set of ordinals. The argument of the Burali-Forti paradox shows that the purported theorem must in fact be false. The order type of the set of ordinals below α has relative type two higher than the type of α; it can be shown by induction to be equivalent to the order type of the double singleton image of an element of α. We define $\mathrm{T}\{\alpha\}$ as the order type of the singleton image of an element of α, by analogy with the type-raising operation already defined on the cardinals: we can show that the order type of the ordinals below α is $\mathrm{T}^2\{\alpha\}$, and the argument of the Burali-Forti paradox shows that $\mathrm{T}^2\{\Omega\} < \Omega$; the order type of the ordinals is smaller than the largest ordinals (there is, of course, no largest ordinal). Now observe that there is an external descending sequence of ordinals $\Omega > \mathrm{T}^2\{\Omega\} > \mathrm{T}^4\{\Omega\} \ldots$; it is easy to see that the T operation respects order. This shows that in a model of *NFU* the "ordinals" cannot be well-ordered; of course, it should also be noted that in a *model* of *NFU* $\in$ must be a relation, so we should expect a model of the theory to be defective. This rules out the possibility of a Replacement Axiom of the most general form for *NFU*, since there are externally countable "proper classes," but we will introduce an adequate replacement scheme of a more restricted character below. The fixed points of the T operator on ordinals are called "Cantorian" ordinals (ordinals of well-orderings whose domains are Cantorian) and ordinals of well-orderings whose domains are strongly Cantorian are called "strongly Cantorian" ordinals.

An aspect of the definitions of cardinal and ordinal numbers which we have not discussed, but which provides a convenient introduction to the use of the Axiom of Choice in *NFU*, is the fact that there is no function which takes A to $|A|$ or a well-ordering to the associated ordinal number. In general, if we take an equivalence relation R, the equivalence class $[x] = \{y \mid yRx\}$ of an element x of the domain of R would be one type higher than x in TT, so we cannot define a map which takes each x to $[x]$. It is thus advantageous to choose representative elements of each equivalence class $[x]$, which would be at the same type as x. The ability to do this is supplied by

Axiom 18. For any disjoint set P of nonempty sets, there is a set $C \subseteq \bigcup P$ whose intersection with each element of P is a singleton.

This is the form of the Axiom of Choice most convenient for this application. The usual equivalences of forms of the Axiom of Choice can be proven in *NFU*. In particular, Axiom 18 implies that there is a well-ordering of the universe, and also that the cardinal numbers are well-ordered by the natural order. In *NF*, where the Axiom of Choice is refutable, the technical inconvenience of being unable to define the map which takes x to $[x]$ in equivalence class constructions is unavoidable. We could use Axiom 18 to redefine cardinal numbers as typical sets of each cardinality and ordinal numbers as typical well-orderings of each order type, but we prefer to stay with the natural definitions in these cases. In other kinds of equivalence class constructions, the availability of Choice can prove to be a useful tool.

A very interesting application of the Axiom of Choice in *NFU* is the ability to prove that there are urelements (which is equivalent to the ability to refute the Axiom of Choice in *NF*). We will prove this using the further assumption of the Axiom of Counting. The original proof of Specker is similar in spirit but somewhat more complicated (a proof of the negation of Choice in *NF* without Counting will be given below). Our proof is adequate for any version of *NFU* in which the natural numbers are standard (since this clearly implies Counting).

We define exp as the function which sends each cardinal κ to 2^{κ}. We define $\mathrm{ST}\{\kappa\}$ for any cardinal κ as the collection of cardinals λ such that $\exp^n(\lambda) = \kappa$ for some natural number n. We consider the set $\mathrm{ST}\{|V|\}$. Let μ be the smallest cardinal in $\mathrm{ST}\{|V|\}$. Consider the sequence s and number n such that $s_i = \exp^i(\mu)$ for $0 \leq i \leq n$, with $\exp^n(\mu) = |V|$. Now consider the sequence $\mathrm{T}s$ of iterated images of $\mathrm{T}\{\mu\}$ under exp: $\mathrm{T}s_i = \exp^i(\mathrm{T}\{\mu\}) = \exp^{T\{i\}}(\mathrm{T}\{\mu\}) = \mathrm{T}\{\exp^i(\mu)\} = \mathrm{T}\{s_i\}$, when s_i is defined (there may be further terms of $\mathrm{T}s$). Note the application of the Axiom of Counting. If $\mathrm{T}\{\mu\}$ is not in $\mathrm{ST}\{|V|\}$, then neither is $\mathrm{T}s_{n+1} = \exp(\mathrm{T}s_n) = \exp(\mathrm{T}\{|V|\}) = \exp(|\mathcal{P}_1\{V\}|) = |\mathcal{P}\{V\}|$, so $\mathcal{P}\{V\} \neq V$, and there are urelements. If $\mathrm{T}\{\mu\}$ is in $\mathrm{ST}\{|V|\}$, then it is greater than or equal to μ by choice of μ, and an easy induction shows that $\mathrm{T}s_i \geq s_i$ for each i for which the latter is defined, so, in particular, $|\mathcal{P}_1\{V\}| = \mathrm{T}s_n \geq s_n = |V|$, which is absurd. Note that we have shown not only that there are urelements, but that most objects in the universe are urelements, since $|\mathcal{P}\{V\}| < |V|$.

The T operation is an external homomorphism from cardinals or ordinals into cardinals or ordinals, respectively, with respect to the usual operations on cardinals and ordinals. This should remind us of the fact that we used external automorphisms of models to construct models of *NFU* above. It turns out that the T operations inside the theory are closely related to the isomorphism j used in constructing models of the theory. We will develop the theory of well-founded extensional relations below, in which it proves possible to duplicate our technique for constructing models of *NFU* in *NFU* itself, using the T operation on isomorphism types of such relations as the external automorphism (since T is not a function in *NFU*, we get an interpretation rather than a model).

At this point, we have developed a theory similar to that in Rosser's *Logic for Mathematicians*. The differences are that urelements are allowed, that the ordered pair is primitive (the theorem above shows that most pairs must be urelements, so there is no hope of a set-theoretic definition of the pair!), and that we provide the full Axiom of Choice. Rosser used the denumerable Axiom of Choice, due to general misgivings about choice; this means that the system of his book (*NF* + Counting + denumerable Choice) may well be consistent, as far as anyone knows. Everything in Rosser's book can be adapted to the system given here, except the definition he gives for the type-level ordered pair.

We reiterate Rosser's claim that this set theory is adequate as a foundation for mathematics. As it stands, it is stronger than the theory of types (Axiom 17 adds considerable strength), but weaker than *ZFC*. We will see below how to extend it so that it is at least as strong as *ZFC*, but the strength of *TT* alone is more than adequate for most classical mathematics.

We claim, moreover, that this set theory is a convenient medium for mathematics. Stratification requires us to apply the singleton operation, one of the T operations, or a similar type-raising device occasionally, but the appropriate use of these is not hard to learn. Mathematical practice is usually stratified in any case, except for confusion of objects with their singletons; this is true even of set theory if the objects of *ZFC* are interpreted as isomorphism types of well-founded extensional relations. We gain the mild technical convenience that many large collections become sets; for example, most natural categories are sets in *NFU*. We will see below that it is possible to arrange for our universe to contain an "inner model" (it will not be a set) of a set theory of the usual type, so Quine-style theories can be understood as strictly extending Zermelo-style theories.

5 Interpreting Set Theory of the Usual Type in *NFU*; the Axiom of Small Ordinals

It is possible to interpret bounded Zermelo set theory in a suitable extension of *NFU* in a natural way. The cumulative hierarchy appears in *NFU* as (part of) the set of isomorphism types of "connected" well-founded extensional relations. A relation R is well-founded if for each subset S of the union of the domain and converse domain of R there is a "minimal" element $s \in S$ such that $\sim rRs$ for all $r \in S$. A relation R is "extensional" if the preimage of an element of the converse domain of R determines the element precisely: if it is the case for every s that sRt if and only if sRu, then $t = u$. A well-founded extensional relation is "connected" if and only if there is an element t of the converse domain of R such that for every s in domain of R there is a finite sequence s_i such that $s_0 = s$, $s_i R s_{i+1}$ whenever the terms are defined, and the last term of the sequence is t. The element t (necessarily unique — a consequence of well-foundedness) is called the "top" of the relation. With each element x in the converse domain of a "connected" well-founded extensional relation R, we can associate a subrelation of R, the maximal connected well-founded extensional subrelation of which it is the "top," called the "component" associated with x. In the usual set theory, each nonempty set is associated with a unique "connected" well-founded extensional relation, that of membership on its transitive closure (understood as including the set itself); the elements of the set have as their associated relations the components associated with the preimages of the "top" of the associated relation of the original set. The empty set presents a technical problem. We handle this by adding the top t to the structure: let (t, R) be a "connected well-founded extensional relation with top" in case R is a "connected well-founded extensional relation" as defined above and t is its "top," or in case R is empty. This allows us to associate components with minimal elements of R as well (elements of the domain which are not elements of the converse domain). We will usually refer to a structure (t, R) simply as R, which is entirely safe if R is nonempty. We will call the components associated with the preimages of the top of a connected well-founded extensional relation the "immediate" components of the relation. We call the elements of the domain and converse domain of a connected well-founded extensional relation the "nodes" of the relation.

We call the set of isomorphism types of connected well-founded extensional relations Z. We define a relation E on Z: xEy holds if an element of x is isomorphic to an immediate component of an element of y. Notice that the type of the transitive closure of a set A (in the usual set theory) stands in the relation E to the type of the transitive closure of a set B exactly when A is an element of B. It is easy to prove that E is a well-founded extensional relation. It is not connected, but this can be corrected; we can form a relation E^+ by adding a top element t such that xE^+t for all $x \in Z$. Now E^+ belongs to some isomorphism type $v \in Z$. For any $x \in Z$ and $R \in x$, we define $\mathrm{T}\{x\}$ as the isomorphism type of the singleton image $\mathrm{SI}\{R\}$ (which is easily seen to be in Z as well, since $\mathrm{SI}\{R\}$ is clearly connected, well-founded, and extensional precisely when R is). Note that $\mathrm{T}\{x\}\ E\ \mathrm{T}\{y\}$ exactly when xEy. The restriction of the relation E to the isomorphism types of the components associated with the nodes of a connected well-founded extensional relation R is parallel in structure to the relation R itself (it is itself the component of E below the type of R); it is straightforward to show that if the isomorphism type of R is x, the isomorphism type of the relation E restricted to isomorphism types of components of R is $\mathrm{T}^2\{x\}$ (isomorphism types of components of R are two types higher than elements of the domain and converse domain of R). We apply this argument to the isomorphism type v of E^+; we see that the isomorphism type of the component of E determined by v itself is $\mathrm{T}^2\{v\}$. It is easy to see that there are connected well-founded extensional relations which have types not in this component; consider the relation obtained by adding a new top node to E^+; it follows that $\mathrm{T}^2\{v\}$ is distinct from v, since a well-founded relation cannot be isomorphic to a proper component of itself. This argument is analogous to the argument from the Burali-Forti paradox that $\mathrm{T}^2\{\Omega\} < \Omega$.

The relation E is being considered as a "membership" relation; thus, we would like to know what "comprehension" principles it satisfies. For what subsets S of Z are there elements x of Z such that for all y, yEx if and only if $y \in S$? This cannot hold for every set S; an obvious counterexample is the set of all $x \in Z$ such that $\sim xEx$, which is actually the set Z itself. But it does hold for every subset S of $\mathrm{T}[Z] = \{\mathrm{T}\{z\} \mid z \in Z\}$. We want to build a well-founded extensional relation which has immediate components in exactly the isomorphism classes in S. The trick is that we can construct elements of each class $\mathrm{T}\{x\}$ in S by replacing each object $\{p\}$ in a chosen element $\mathrm{SI}\{R\}$ of $\mathrm{T}\{x\}$ with $(\{p\}, R)$; these relations are disjoint. Take the union of this collection of relations and add a top element related to the top of each of the subrelations, and we get an element of the desired isomorphism class.

Note the essential use of the Axiom of Choice in constructing relation types representing subsets of $\mathrm{T}[Z]$; in *NF* it would not be possible to choose a single element of each relation type as we did above. Without the Axiom of Choice, we could construct a connected well-founded relation which would not be extensional by taking disjoint copies of every relation in each class in $\mathrm{T}[S]$ and adding a top node, but we would have no way to pick *one* relation from each class. There is a way around this. It is possible to choose a canonical element of each class $\mathrm{T}^2\{x\}$, namely, E restricted to the component of E determined by x; any collection of such canonical elements of classes can be made disjoint by pairing each node of each relation with its top node. Thus, in the absence of the Axiom of Choice, as in *NF*, we can find relation types representing each subset of $\mathrm{T}^2[Z]$.

If we let S be $\mathrm{T}[Z]$ itself, we get a relation R in a class x which has components associated with preimages of its top node in each class in $\mathrm{T}[Z]$. The paradoxical thing about this is that one of these components must be of type $\mathrm{T}\{x\}$, and must in its turn have a component of type $\mathrm{T}^2\{x\}$, and so forth. The top nodes of these components would *seem* to form a descending chain in the relation R, which is impossible, since R is well-founded! But the definition of the descending chain is unstratified, so no contradiction ensues. This is analogous to the result that there is an external descending sequence in the ordinals.

We now present the local version of the cumulative hierarchy of the usual set theory. If S is a subset of Z, let $P(S)$ be the collection of elements of Z whose preimages under E belong to S; if the collection of preimages under E of elements of $P(S)$ is $\mathcal{P}\{S\}$, we call S *complete*. It should be clear that $|P(S)| =$

$\exp(|S|)$ if S is complete. Let H be the intersection of all sets of subsets of Z which are closed under the P operation (which is a function) and contain the unions of all chains of their elements under inclusion. H is an analogue of the cumulative hierarchy of the usual set theory; we call an element of H a "stage". The collection of "transitive" subsets of Z (subsets S such that $S \subseteq P(S)$) is closed under the P operation and closed under union of nested chains of its elements, so all elements of H are transitive. We can then prove similarly that the elements of H below any stage x are well-ordered by inclusion, with successors being images under P and unions being taken at limits. Consider the set of stages x for which this is true. It should be clear that $\bigcup H = P(\bigcup H)$, and that $\bigcup H = Z$.

Each ordinal belongs to Z (well-orderings are connected well-founded extensional relations!), and so belongs to some stage, and to a first stage. It should be clear that the first stage to which $\alpha + 1$ belongs is the successor of the first stage to which α belongs, and that the first stage to which the limit of a sequence of ordinals belongs is the limit of the stages at which the ordinals in the sequence first appear. No stage can contain more than one new ordinal; but not every stage can contain even one new ordinal: if the last stage Z contained a new ordinal α, the ordinal $\alpha + 1$ would not belong to any stage.

The image under T (T^2 in the absence of AC) of a stage with no incomplete predecessors is clearly a complete stage. The image under T (T^2 in the absence of AC, as before) of a stage with incomplete predecessors would not be a stage.

Now consider the complete stages. Z is not complete. There is a first incomplete stage Z_0; it might be Z itself. Since the image under T (T^2 in the absence of AC) of Z_0 is complete and a stage, it is distinct from Z_0 (and so below Z_0). The cardinalities of the complete stages are the beth numbers. Each complete successor stage must contain a new ordinal; we can define the ordinal rank of any stage at or below Z_0 as the new ordinal appearing in the successor stage to its image under T (note the type differential; T^2 would be used in the absence of Choice) "realizing" the set of ordinals appearing in previous stages. Interesting stages to consider are the first incomplete stage (whose ordinal rank is the rank of the initial segment of the cumulative hierarchy realized in our system) and the first stage which does not contain a new ordinal (not necessarily the same).

We can now present an interpretation of *NFU* inside *NFU* which is precisely analogous to the construction of a model of *NFU* in the usual set theory given above. Recall that Z_0 is defined as the first incomplete stage (the initial segment of the cumulative hierarchy encoded in our system). The elements of Z_0 will be the objects of our interpretation. We define "$x \in' y$" for $x, y \in Z_0$ as "$T\{x\}\ E\ y$ and $y\ E\ P(\mathrm{T}[Z_0])$". We omit the proof that this is an interpretation of *NFU*, but it is easy to reconstruct if one bears in mind that E is analogous to membership in the underlying nonstandard model of set theory in the original construction, T is analogous to the external automorphism j (actually, analogous to its inverse), Z_0 is analogous to V_α, and $P(\mathrm{T}[Z_0])$ is analogous to $V_{j^{-1}(\alpha)+1}$. The slight technical problem is that one does not have access to analogues of stages above stage α of the cumulative hierarchy in the original construction, but this can be remedied by considering the fact that j projects these levels to lower levels. A version of this interpretation could be carried out as well in *NF* (in which the equivalence relation interpreting equality must be nontrivial; see [14] for details), but it would yield an interpretation of *NFU* with urelements in *NF*, not an interpretation of *NF*; elements of the analogue of $Z_0 - P(\mathrm{T}[Z_0])$ will be interpreted as urelements. An interpretation is constructed here rather than a model, because $\in'$ is not a relation; its definition is unstratified (with the same type differential as in the case of $\in$ itself).

We observe that the models of *NFU* described above and the interpretation of *NFU* just outlined satisfy an additional axiom not discussed so far. In the models described initially, consider the map which sends the singleton of x to x for each x in $V_{j(\alpha)}$. This map will be interpreted in the model of *NFU* as sending the singleton of x to $j^{-1}(x)$; it implements the external map j^{-1} as a type-raising operation! Note that $j^{-1}(x)$ is always interpreted as a set. The internal interpretation of *NFU* has an analogous map, the one which sends the type of a relation whose top node has one preimage under E to the type of the relation obtained by omitting that top node, restricted to the case where the relation with top node removed is in $T[Z]$. The axiom satisfied in both cases is the

Axiom of Endomorphism: There is a one-to-one map Endo from $\mathcal{P}_1\{V\}$ into $\mathcal{P}\{V\}$ (a map from singletons into sets) such that for each set A, $\mathrm{Endo}(\{A\}) = \{\mathrm{Endo}(\{B\}) \mid B \in A\}$.

The fact that Endo implements j^{-1} in models makes it easy to verify this. Note that an urelement u has $\mathrm{Endo}(\{u\})$ a set as well, but we have less control over what this set may be. We will write $\mathrm{Endo}\{A\}$ instead of $\mathrm{Endo}(\{A\})$ hereinafter. This axiom is mostly of technical interest, and we will not officially adopt it as an axiom of our system; note that it adds no strength, since it holds in the interpretation of *NFU* on Z_0. It is interesting to note that general "accessible pointed graphs" can be used to develop an interpretation of a Zermelo-style set theory with Aczel's Axiom of Anti-Foundation (*AFA*) (see [1]), which will generate an internal interpretation of *NFU* inside *NFU* which satisfies Endomorphism and an axiom related to *AFA* which implies *AFA* for strongly Cantorian graphs. This is developed in our paper [14]. The Axiom of Endomorphism is false in *NF* as we showed in [14] (and so implies the existence of urelements, but it also gives structure to urelements, by associating each urelement u with the set $\mathrm{Endo}\{u\}$).

Another tempting candidate for an axiom, which holds in sensible models, is "$Z = Z_0$".

We observe that the stages of Z with strongly Cantorian rank provide an interpretation of bounded Zermelo set theory with E as membership. Clearly each strongly Cantorian rank is complete (every ordinal below a strongly Cantorian ordinal is strongly Cantorian, and the rank of the first incomplete stage Z_0 is moved by T). Bounded Zermelo set theory is interpreted because set definitions involving quantification over all strongly Cantorian stages cannot be encoded, since the strongly Cantorian ordinals, and so the union of the corresponding stages, do not make up a set. We are trying to capture the idea that the notion of "strongly Cantorian" captures the size of the sets of the usual kind of set theory. A warning is needed: Axiom 17 is needed to establish Infinity in this interpretation of bounded Zermelo set theory; otherwise, all strongly Cantorian ordinals might be finite.

It is convenient, as it is in *ZFC*, to be able to talk about proper classes. Quine introduced the use of proper classes in *NF* in his book *Mathematical Logic* ([25]); the system of this book is commonly known as *ML*. The original system of *ML* was inconsistent; one should be careful that one has the later editions which correct the original error. *ML* adds an impredicative system of proper classes to *NF*, and is stronger than *NF*, although Wang has shown that the corrected *ML* is consistent if *NF* is consistent. Our use of classes here will be restricted to the sort which can be reduced to talk of formulas, and so involves no strengthening of our theory.

It is possible to arrange for the relation E to coincide with membership $\in$ on the strongly Cantorian stages by using a permutation technique to redefine membership. If π is a permutation of V, we can redefine membership, defining "$x \in' y$" as "$x \in \pi(y)$," and the resulting system will still satisfy the axioms of *NFU*. The application of this method to *NF* is due to Dana Scott in [29]; all stratified sentences are invariant under such permutations, and C. W. Henson showed in [18] that some unstratified sentences, such as the Axiom of Counting, are also invariant. Forster shows in [9] that, if one allows certain external permutations, the class of invariant sentences can be made to be precisely the class of stratified sentences. Such permutation techniques are almost the only ones known for obtaining consistency and independence results relative to *NF* (the methods known for the usual set theory can be adapted to *NFU*). The permutation π which we will use interchanges each element $\mathrm{T}\{x\}$ of $\mathrm{T}[Z]$ with the set of preimages of x under E; the application of T is needed for stratification. Each erstwhile element of $\mathrm{T}[Z]$ gets the preimages under E of its preimage under T as its new elements; strongly Cantorian elements of Z get their own preimages under E as their new elements. Thus, we have a class inner model of bounded Zermelo set theory in any model of *NFU* + Counting.

We introduce a notion of "small" ordinal; an ordinal is "small" if it is less than some Cantorian ordinal. The notion of "strongly Cantorian" as a notion of smallness needs no corresponding refinement, since an ordinal less than a strongly Cantorian ordinal is necessarily strongly Cantorian. A more powerful notion of "smallness," with which the following development could be duplicated, asserts that an ordinal α is small if it is less than some ordinal β such that $\beta \leq \mathrm{T}\{\beta\}$.

We introduce a powerful axiom scheme to be adjoined to the theory of the previous section which has the effect that the notions of "smallness" so far proposed (Cantorian, strongly Cantorian, and "small" itself) are collapsed together and the stages of Z of small ordinal rank come to interpret full *ZFC* (at least). With the application of the permutation technique described above, we get a class inner model of *ZFC* in our theory. We call it the Axiom (Scheme) of Small Ordinals.

Axiom scheme 19. For any formula ϕ (not necessarily stratified) there is a set B whose small ordinal elements are exactly the small ordinals x such that ϕ.

The strength of our full theory with axiom scheme 19 is at least that of *ZFC*; we conjecture that this is exactly its strength, but we do not know how to prove this. We know how to model the full theory in the presence of a measurable cardinal (using elementary embeddings — to be discussed in our pending [17]). The motivation of the Axiom of Small Ordinals is that the small ordinals are to be identified as the standard ordinals and the set of all ordinals in *NFU* is to be understood as an initial segment of a nonstandard extension of the ordinals. The intuitive argument for the Axiom is that the condition ϕ, whether stratified or not, does define a class of standard ordinals, and the analogue of that standard class in the nonstandard extension of the ordinals, intersected with the set of ordinals of *NFU*, will give the desired set B. We do not see how to implement this except with large cardinal axioms, but our experience with the Axiom leads us to believe that it gives no more strength than *ZFC*. A merit of this axiom is that it is natural in terms of *NFU*; it may be considered as an extension of the possibilities of subversion of stratification restrictions on set formation when variables are restricted to strongly Cantorian sets. Axioms like "there is an inaccessible cardinal" (which implies that there is a Cantorian inaccessible cardinal), which certainly give at least the strength of *ZFC* (if an inner model of *ZFC* is desired, the inaccessible cardinal should be strongly Cantorian) have two vices: their motivation is extrinsic to *NFU*, and they are generally too strong; the axiom above, for instance, defines a theory stronger than *ZFC*. It is not easy to define a natural extension of *NFU* which is precisely equiconsistent with *ZFC*, as we hope that we have done with axiom scheme 19. Even if the Axiom of Small Ordinals is too strong (it might be a large cardinal axiom) it remains a natural axiom with which to extend *NFU*. In class language, the Axiom asserts that each class of small ordinals is the intersection of a set with the class of all small ordinals.

We present applications of the Axiom of Small Ordinals. First of all, axiom 17 is now redundant. We can prove the Axiom of Counting as follows: there is a set B whose small ordinal members are exactly the finite ordinals n such that $\mathrm{T}\{n\} = n$. Consider the smallest ordinal m not in this set. If it is a finite ordinal, then $\mathrm{T}\{m-1\} = m-1$ and $\mathrm{T}\{m\} \neq m$, which is absurd. It must then be ω, and it is established that $\mathrm{T}\{n\} = n$ for each finite ordinal n, and so the analogous result holds for finite cardinal numbers as well.

More generally, we can prove that every Cantorian ordinal is strongly Cantorian, which implies that every small ordinal is strongly Cantorian, and these three notions collapse together. Let α be a Cantorian ordinal. There is a set B whose small ordinal elements are exactly the strongly Cantorian ordinals less than α; the complement of this set has a smallest element β. It is straightforward to prove that successors of strongly Cantorian ordinals are strongly Cantorian and limits of sets of strongly Cantorian ordinals are strongly Cantorian; it follows that β is strongly Cantorian and equal to α. We have only one notion of "smallness," for which we will use the terms "small" and "Cantorian" henceforth.

We have already shown that a Replacement scheme of the most obvious form is impossible in *NFU*, since there are externally countable proper classes. But we *can* prove that any class of Cantorian ordinals which can be placed in an external one-to-one correspondence with a Cantorian set is a set (necessarily Cantorian). The proof goes as follows: code pairs of ordinals by ordinals in a natural way (pairs of Cantorian ordinals should correspond exactly to the Cantorian ordinals, which will hold under the usual scheme). Suppose that for each element x of a Cantorian set A there is exactly one Cantorian ordinal y such that ϕ. By the Axiom of Choice, we may suppose without loss of generality that A is the set of ordinals less than some Cantorian ordinal α (a bijection exists between any Cantorian A and such a segment of the ordinals). Consider the class of ordinals which code pairs of ordinals (x, y) such that $x < \alpha$ and ϕ; such ordinals are small, and by the Axiom of Small Ordinals there is a set B whose small ordinal elements are exactly the elements of this class. Now consider the set $C = \{y \mid \text{for some } x < \alpha,\ y \text{ is the smallest ordinal such that } (x, y) \in B\}$; this is the desired set. A version of Separation is easier to prove yet: let α be a Cantorian ordinal; then $\{x < \alpha \mid \phi\}$ is the intersection of a set B with the class of small ordinals, and the intersection of B with the set of ordinals less than α is the desired set. Another useful result provable with the Axiom of Small Ordinals is that any set A of Cantorian ordinals is a Cantorian set: consider the strictly monotone map from the set A to an initial segment of the ordinals; if its range contains a non-Cantorian ordinal, there is a non-Cantorian ordinal less than a (strongly) Cantorian ordinal, which is impossible. A similar argument shows that any class of Cantorian ordinals is either a Cantorian set or is placed in one-to-one correspondence with the class of Cantorian ordinals by the strictly monotone class map from the class onto an initial (class) segment of the ordinals.

It is easy to show that $\exp(\kappa)$ is Cantorian if κ is Cantorian in *NFU* alone. This observation allows us to encode the small stages of the cumulative hierarchy into the small ordinals: when V_β has been encoded into the ordinals of well-orderings of sets with cardinality less than $|V_\beta|$ for each ordinal β less than an ordinal α, V_α has already been encoded if α is limit; if α is successor, encode the sets in $V_\alpha - V_{\alpha-1}$ into the ordinals of well-orderings of sets with cardinality greater than or equal to $|V_{\alpha-1}|$ and strictly less than $|V_\alpha|$, a collection of ordinals which is easily seen to be the right size. This process can be generalized to an encoding of elements of Z by general ordinals, but some applications of type-raising operations would be needed to recover stratification; the elements of the stages with small index would be encoded by exactly the small ordinals. Use of this encoding enables us to apply Separation and Replacement without the restriction to bounded quantifiers in building sets in the interpretation of the cumulative hierarchy, which gives us the full strength of *ZFC* (technical details are omitted here). If the Axiom of Choice is not desired, it is sufficient for interpreting *ZF* to have a stronger version of axiom scheme 19 which applies to elements of Z rather than ordinals.

We hope that we have now indicated that *NFU*, suitably extended, is an adequate medium for doing mathematics. The habit of maintaining stratification is not hard to learn, and the stratification restrictions can be evaded in many applications by using the Axiom of Counting or the Axiom of Small Ordinals. In return, one recovers attractive definitions of cardinal and ordinal number, and most of the constructions of naive set theory. The notion of a set of the cumulative hierarchy is presented as a generalization of the notion of ordinal number. It is interesting to observe that the surreal numbers of Conway (see [3]) can be constructed as equivalence classes of suitable relations (his construction is analogous to the von Neumann definition of the ordinals, and so is unstratified), and that the surreal numbers form a set in *NFU*. Most interesting categories are sets in *NFU*, such as the category of sets, the category of topological spaces, the category of groups or the category of categories itself (though it should be noted that the properties of these big categories are not quite what one might expect; see [21]). A text which develops the basics in the system of our axioms 1 through 18 already exists in Rosser's [27] (read in accordance with our instructions above).

6 What is Wrong with *NF*?; "Tangled" Type Theories Introduced

In all of this, we have said little about *NF* itself, except to indicate that *NF* + Counting proves that AC is false (since *NFU* + Counting + Choice proves that there are urelements), and to comment where the lack of Choice would affect constructions carried out above if they were done in *NF*. We have seen that the internal interpretation of *NFU* in *NFU* would give an interpretation of *NFU* (not of *NF*!) if carried out in *NF*.

We will not give Specker's original refutation of the Axiom of Choice in *NF* here, but we give an alternative proof. Suppose that the Axiom of Choice holds. Consider the cardinalities of the complete stages in Z (the beth numbers).

Define $\beth_\alpha$ as the cardinality of the stage with ordinal rank α, when all stages below stage α are complete. Note that $\beth_{\alpha+1} = \exp(\beth_\alpha)$ for all α in the domain of definition. Consider the first cardinal of an incomplete stage, $\beth_\zeta$, which is the last beth number. Clearly, $\beth_\zeta > \mathrm{T}\{|V|\}$, as otherwise we could construct another complete stage (without choice, these cardinals might be incomparable). $\beth_\zeta$ must be the first beth number greater than $\mathrm{T}\{|V|\}$, unless $\beth_{\zeta-1}$ is the first and $\beth_\zeta = |V|$, because $\exp(\mathrm{T}\{|V|\}) = \exp(|\mathcal{P}_1\{V\}| = |\mathcal{P}\{V\}|$, which is equal to $|V|$ because we are in *NF* (it follows also that $\exp(\mathrm{T}^{n+1}\{|V|\}) = \mathrm{T}^n\{|V|\}$ for each concrete n). In either case, there is a first beth number greater than or equal to $\mathrm{T}\{|V|\}$; call it $\beth_{\zeta_0}$. Now the first beth number greater than or equal to $\mathrm{T}^2\{|V|\}$ must be $\beth_{T\{\zeta_0\}}$ and must be less than $\mathrm{T}\{|V|\}$, and the next beth number after this must be greater than or equal to $\mathrm{T}\{|V|\}$, and so must be $\beth_{\zeta_0}$ itself. But this implies that $\mathrm{T}\{\zeta_0\} = \zeta_0 - 1$, which is impossible: consider the parity (even or odd) of the finite part of the ordinal ζ_0, which must be preserved by the T operation (the finite parts of ζ_0 and $\mathrm{T}\{\zeta_0\}$ need not be equal themselves, since we are not assuming the Axiom of Counting, but their parity will be the same). The contradiction establishes that the Axiom of Choice is false in *NF*.

This argument shows that there cannot be a first beth number greater than or equal to $\mathrm{T}\{|V|\}$, and so, since the last beth number is not less than or equal to $\mathrm{T}\{|V|\}$, it must be incomparable with $\mathrm{T}\{|V|\}$.

An important tool in the study of big cardinals (we use this term to denote cardinals near the cardinality of the universe, as opposed to large cardinals in the usual sense) in *NF* is the "Specker tree" $\mathrm{ST}\{\kappa\}$ (already used in our proof of the existence of urelements in the system of axioms 1 through 18 above), the set of iterated preimages of a cardinal κ under exp, with the cardinal in question being the top of the tree and the nodes immediately below any node of the tree being its preimages under exp. A classical theorem of Sierpinski, provable in *NF*, asserts that $\aleph(\kappa) < \exp^2(\kappa)$, where $\aleph(\kappa)$ is the first cardinal of a well-ordered set not less than or equal to κ, or the image (if any) under T^{-2} of the cardinality of the set of order types of well-ordered subsets of a set of cardinality κ ($\aleph$ is a stratified version of Hartog's function). We have improved the exponent in Sierpinski's original result by using a type-level formulation of the ordered pair due to Quine (see [27]) in place of the usual ordered pair, which raises types by two; this is a minor technical change which makes no essential difference in the sequel. We prove that the Specker tree of any cardinal is well-founded: suppose that a subset S of the tree had no minimal element (element with no preimage in exp also in S). Consider the smallest cardinal $\aleph(\kappa)$ for κ in the Specker tree; choose a κ realizing this value, and consider a preimage λ of κ under $\exp^2$ in S, which exists by hypothesis; by the result of Sierpinski cited above, $\aleph(\lambda) < \exp^2(\lambda) = \kappa \not\geq \aleph(\kappa)$, contradicting minimality of $\aleph(\kappa)$. Note that the concept of Specker trees makes sense in *ZF* as in *NF(U)*, and that the theorem that Specker trees are well-founded is valid there as well. This theorem is due to Forster, and the proof can be found in [9].

ST$\{|V|\}$, the Specker tree of the cardinality of the universe, contains the "descending sequence" of iterated preimages of $|V|$ under T, since this is also a "descending sequence" of iterated images of $|V|$ under exp. Since Specker trees are well-founded, this "sequence" cannot be a set. To avoid sethood for this sequence, ST$\{|V|\}$ must be "bushy"; there must be many descending paths in ST$\{|V|\}$, because if there were only one, it would be the external sequence and could be defined as a set (a parity argument is required to complete this argument if not all natural numbers are standard). Moreover, the same argument can be applied to any subtree of the Specker tree which contains all elements of the external sequence. ST$\{|V|\}$ also has the property of having its own image under T as an immediate component (ST$\{$T$\{|V|\}\}$). As we have seen above, the strange properties of ST$\{|V|\}$ enable us to disprove Choice in *NF* + Counting.

The bizarre properties of *NF* as opposed to *NFU* are superficially rather surprising; on the face of it, it is unclear why the assumption of strong extensionality has such profound effects. It should be clear, though, that the sticking point is the fact that exp sends T$\{|V|\}$ to $|V|$ if $\mathcal{P}\{V\} = V$, so introducing the external descending sequence in the Specker tree of $|V|$. We believe that we can explain what is wrong with *NF*; we will adapt Jensen's proof of Con(*NFU*) to produce a proof of the equiconsistency of *NF* with a version of type theory whose strangeness should be obvious. Moreover, we are able to convert this type theory to a Zermelo-style theory, still equiconsistent with *NF*, in which the "obvious" strangeness which *NF* shares with *NFU*, the presence of big sets like the universe and the set of all ordinals, disappears. The reasons for the presence of "bushy" Specker trees and the failure of the Axiom of Choice will be made clearer.

The crucial feature of the version of TTU interpreted in the proofs of Con(*NFU*) given above was that it was possible to define a membership relation $\in_{\alpha,\beta}$ of type α objects (members of V_α) in type β objects (members of V_β) for any $\alpha < \beta$, in contrast to the situation in the usual type theory, where β would have to be the successor of α. This was easy to implement in TTU, since we could interpret the excess objects in $V_\beta - V_{\alpha+1}$ as urelements. It is unclear why one would think that one could do this in extensional TT. But we now show that the consistency of *NF* is precisely equivalent to the question of whether such a version of TT is consistent.

Let "tangled type theory" (TTT) be a first-order many-sorted theory with sorts indexed by the non-negative integers (or, more generally, any linearly ordered set with no upper bound), with equality and membership. An atomic sentence "$x \in y$" will be well-formed exactly when the type of x is less than the type of y; an atomic sentence "$x = y$" will be well-formed exactly when x and y have the same type.

If ϕ is a well-formed formula of TT and s is a strictly increasing function from the non-negative integers into the set of types (sequence of types with non-negative indices), the formula ϕ^s produced by replacing each type index i in ϕ with s_i is a well-formed formula of TTT. The axioms of TTT will be exactly the formulae ϕ^s for s any appropriate sequence and ϕ an axiom of TT.

Each object of type τ will have independent interpretations as a set of type σ objects for each $\sigma < \tau$; each of the extensions of an object will determine it uniquely. Relative to any ascending sequence of types, it will be possible to construct sets exactly as they are constructed in TT.

We claim that TTT is consistent exactly if *NF* is consistent. If *NF* is consistent, let M be the universe of a model of *NF*. Let $M \times \{i\}$ be used as type i for each i, and let "$(x, i) \in_{TTT} (y, j)$" be defined as "$x \in_{NF} y$," where $\in_{NF}$ is the membership relation of the model of *NF*; it should be clear that this gives us a model of TTT with $\in_{TTT}$ as the membership relation.

Suppose that we have a model of TTT. Let Σ be a finite set of formulas of the language of TT. Let n be the number of types which occur in Σ. Let A be any n-element set of types; we define ϕ^A as ϕ^s, where s is a strictly increasing sequence of types whose first n elements are the elements of A. Now Σ and our model of TTT determine a partition of the collection of n-element sets of types into finitely many parts, with the element of the partition into which a set A falls being determined by the truth-values of ϕ^A for each ϕ in Σ. This partition has an infinite homogeneous set H by Ramsey's theorem, which is the range of a unique strictly increasing sequence h. The interpretation of TT obtained by interpreting each formula ϕ as ϕ^h satisfies the scheme of typical ambiguity $\phi \iff \phi^+$ for each formula ϕ in Σ. Compactness gives us the consistency of the full scheme of typical ambiguity, and Con(*NF*) follows by the model theory results of Specker cited above.

It should be clear that TTT with any infinite linearly ordered set of types is equiconsistent with TTT with the usual types, by a compactness argument.

We can also construct models of *NF* from models of TTT using automorphisms. Suppose we have a model of TTT; work inside a model of set theory with an external automorphism j. Suppose that $j(\tau) > \tau$ for some type τ. Let type τ in the model of TTT be taken to be the model of *NF*, with membership "$x \in_{NF} y$" defined as "$x \in_{TTT} j(y)$"; note that $j(y)$ is of the higher type $j(\tau)$. The proof that this is a model of *NF* is precisely analogous to the proof that the model of *NFU* given above was a model of *NFU*. If we wanted to model *NF* + Counting in this way, we would need all natural numbers to be fixed by j, so we would need to use more types.

Tangled type theory with urelements proves quite easy to model; it is quite unclear how to model TTT with its strong extensionality.

It is possible in TT to model part of the type structure internally. If we work with type n as the universe, we can use the singletons of type $n-1$ objects to represent type $n-1$ objects, the set of double singletons of type $n-2$ objects to represent type $n-2$ objects, and so forth (notice that the sets representing types are actually found in type $n+1$). Inclusion with its domain (but not its range!) restricted to the set of singletons then represents membership of type $n-1$ objects in type n objects, while the singleton image of this relation represents membership of type $n-2$ objects in type $n-1$ objects. In general, type $n-i$

is represented by the set of i-fold singletons in type n, and the membership relation of type $n-i-1$ in type $n-i$ is represented by the ith iterated singleton image of the inclusion relation restricted to singletons. We can thus represent the theory of n types of TT internally to TT for each concrete n. This technique can be carried out in NF (or NFU) as well. In NF, if V is taken to represent type n, type $n-i$ will be represented by $\mathcal{P}_1^i\{V\}$, and a restriction of $\mathrm{SI}^i\{\subseteq\}$ will represent membership of type $n-i-1$ in type $n-i$. This construction can only be carried out to each concrete finite stage; we have shown above that the set of cardinalities of the types of this model is not a set, so the complete model cannot be a set. We will return to this below when we discuss Orey's proof of the independence of the Axiom of Counting from NF.

In TTT, the same technique can be used to "introspect" on the type structure. But the internal picture of the type structure is quite different from the external one. Choose a type τ and fix a sequence of types above τ in which to define sets of type τ objects, cardinals of such sets, and so forth. Now observe that each sequence of types below τ with our fixed sequence of higher types appended is associated with an interpretation of TT, which can introspect on its own type structure exactly as in the preceding paragraph. The cardinals of each set representing a type in each of these interpretations is mapped to the cardinal of the set interpreting the next higher type in the sequence by the exponential map (as defined for cardinals of sets of type τ objects – the identity of this map as defined in any of these sequences of types depends only on the fixed sequence of types above type τ). This ensures that the different representations of each type $\sigma < \tau$, one for each finite sequence of types beginning with σ and ending with τ, are different, because their cardinals have different behavior under the exponential map. The cardinals of the various representations of types below τ are seen to be elements of the Specker tree of the cardinal of type τ itself (considered as a set in the type above it); so we see that the Specker tree of type τ has the "bushiness" already seen in NF.

Instead of investigating the internal type structure of TTT in detail, we present a new type theory in which the types form a tree. The types will be the ascending sequences of non-negative integers whose ranges are cofinite. If s is a sequence, we define the sequence s^+, called the "tail" of s so that $s^+_i = s_{i+1}$. A sentence "$x \in y$" will be well-formed exactly when the type of y is the tail of the type of x; a sentence "$x = y$" will be well-formed, as usual, exactly when the types of x and y are the same. The axioms of the type theory are those which are obtained from axioms of TT by substituting types of this system for types of TT in such a way as to preserve well-formedness, plus an additional axiom scheme which provides that the truth value of a sentence is completely determined by the first elements s_0 of the types s which occur in it. The motivation is that each sequence s is to be understood as the index of a version of type s_0 in TTT; the additional information specifies what sequence of higher types we are to use, and we recover the uniqueness of the "power set" type for each type. The fact that the sequences used as type indices are increasing and the "power set" of a type is the tail of the sequence which is its index ensures that our sentences translate to sentences of TTT; the additional condition ensures that the truth values of sentences which translate to the same sentence of TTT are the same. The simple coding of the types as sequences given here was suggested privately by Robert Solovay; the author originally used a more complicated way of representing the tree of types. The proof of equiconsistency of this system with TTT is easy.

An advantage of this scheme is that it can be converted into a version of bounded Zermelo set theory (Zermelo set theory with Δ_0 comprehension). Foundation must be weakened to allow "Quine atoms" (objects which are their own sole elements). In bounded Zermelo set theory, each cardinal κ determines an interpretation of TT, in which type i is the ith iterated power set of a set of cardinality κ and membership is defined in the obvious way. The theory of this interpretation of TT is independent of the choice of base set for the model; it depends only on the cardinal κ. Bounded Zermelo theory with Quine atoms and cardinals κ_s for each sequence s which is a type in the theory of the previous paragraph, with axiom schemes asserting that $\exp(\kappa_s) = \kappa_{s^+}$ for each concrete sequence s and that each concrete sentence ϕ in the type theory of i types with a base set of cardinality κ_s has truth value depending only on the first i elements of the sequence s is equiconsistent with the theory of the previous paragraph, and so with TTT and *NF*. The way in which this version of bounded Zermelo set theory interprets the theory of the previous paragraph is obvious; we indicated above how bounded Zermelo set theory can be interpreted in the usual theory of types — use the method indicated there (noting that elements of type 0 become Quine atoms), with type i being intepreted as the ith iterated image of any fixed sequence s' under the tail operation, to interpret this version of bounded Zermelo set theory in the type theory of the previous paragraph. In the interpretation of bounded Zermelo set theory in type theory, the cardinality of any set is identified with the cardinality of its singleton image in the next type up; each type s in the theory of the previous paragraph has an iterated image under tail which is also an iterated image under tail of s'; the cardinal of the iterated image of s under the relevant singleton operations in any type which is also an iterated image under tail of s' will be interpreted as κ_s. These constructions are discussed in detail in our upcoming paper [16].

An advantage of the reduction of the question of the consistency of *NF* to the consistency of such an extension of bounded Zermelo set theory is that it shows that the existence of "big" sets in *NF* has nothing to do with the consistency question; the problematic aspects of *NF* can be expressed in terms of patterns of cardinals in a set theory of the usual type without choice. The features which *NF* shares with *NFU* are not the problem.

The distinctive results about *NF* have been or can be obtained with the aid of the Specker tree of the cardinality of the universe. This construction shows that *all* distinctive features of *NF* can be expected to be derivable from the properties of the Specker tree of the cardinality of the universe, since $|V|$

will correspond to one of the cardinals κ_s in this theory (a nonstandard one), and its Specker tree will include all cardinals corresponding to sequences which have s as an iterated tail (the tail operation on sequences corresponds to the exp operation on cardinals). In particular, the descending sequence of iterated images of $|V|$ under T will correspond to (nonstandard) cardinals of this kind.

The Axiom of Choice is false in sufficiently high types of TTT by an adaptation of Specker's original proof or the one we gave above. But the resulting proof is likely to be very large (because of the application of Ramsey's theorem involved). We give a more natural proof of $\sim$AC in TTT. Like the proof we gave above for *NF*, it relies on the properties of Z_0. Assume that AC holds. In TT, the version of Z_0 in any type can be embedded in the version in the next higher type via a natural type-raising operation. In the presence of AC (also in TT), it is easy to see that going up one type has the effect of adding one or two stages to Z_0 (one or two new beth numbers are added). But going "one type up" in TTT can take us from any given type to any higher type. Work in the sequence of types 0–12. We can first define Z_0 in type 3. The properties of Z_0 at any type in a sequence of types are determined by that type and the three previous types in the sequence. As we go up from type 3 to type 12, the ordinal rank of Z_0 increases by between 9 and 18. Now consider the sequence of types 0–3 followed by 9–12. As we go from type 3 up to type 12 in this sequence of types, the rank of Z_0 goes up by between 4 and 8. This means that the value of the finite part of the two ordinal ranks associated with type 12 must differ mod 19; but this is impossible, since the three types preceding 12 are the same in both sequences, so the ordinal ranks in question must be the same!

There are some interesting fragments of *NF* which satisfy the strong extensionality axiom and are known to be consistent.

The first of these to be discovered was NF_3, whose axioms are extensionality and those instances of stratified comprehension in which it is not necessary to use more than three types. Grishin demonstrated the consistency of NF_3 in [10], and showed that $NF_4 = NF$. Specker's model theory results can be adapted to show that NF_3 is equiconsistent with TT_3 (TT with only 3 types) with the axiom scheme of ambiguity $\phi \iff \phi^+$ (which only makes sense for ϕ mentioning only the lowest 2 types). Now it turns out that all infinite models of TT_3 satisfy the scheme of ambiguity, so NF_3 is a theory with a certain general interest.

Doing mathematics in NF_3 is not very easy, basically because the theory lacks a useful ordered pair. The consistency strength and expressive power of NF_3 with an appropriate form of the Axiom of Infinity have been shown to be precisely the consistency strength and expressive power of second-order arithmetic (by Pabion in [23]), but the coding required to show this is rather devious. Grishin showed that NF_3 + the existence of $\{\{\{x\}, y\} \mid x \in y\}$ is precisely *NF*, so it follows that NF_3 + the existence of a type-level ordered pair = *NF*, because one would be able to define $\{(\{x\}, y) \mid x \in y\}$, and then use this to define the set of Grishin.

In [22], Orey proved interesting results about *NF* by applying the results of Grishin internally to *NF*. As we noted above, we can define a model of TT_4 in *NF* using $\mathcal{P}_1^{3-i}\{V\}$ as type i for i = 0,1,2,3, and using a restriction of inclusion and its first two iterated singleton images as membership relations. It is then possible to define satisfaction for Gödel numbers of formulas in this model in a standard way. The model of TT_4 obtained is ambiguous (satisfies $\phi \iff \phi^+$) for each standard ϕ, but the result which can be proven in *NF* is that $\mathrm{T}\{\phi\} \iff \phi^+$, because the three lower types and their membership operations are the images under type-raising operations of the three upper types and their membership operations. This implies that *NF* + Counting, in which we can draw the further conclusion that $\mathrm{T}\{\phi\} = \phi$, proves Con($TT_4$ + Ambiguity), and so proves Con(NF_4) and Con(*NF*). By Gödel's second incompleteness theorem, the Axiom of Counting must be independent of *NF* if *NF* is consistent. An analogous result can be proven in *NFU*, where it is known to be nontrivial. This is the only independence result we know of for *NF* other than those which can be obtained using permutation methods. It is interesting to observe that *NF* + Counting proves that $\mathrm{ST}\{|V|\}$ has descending paths of each finite length; it is not known whether the existence of Specker trees with descending paths of each finite length is consistent with *ZF* (it is easily shown to be impossible in *ZFC*). Forster has speculated (in [9]) that this may indicate that *NF* + Counting is a very strong theory.

The other extensional fragment of interest is *NFI*, the version of *NF* with strong extensionality and with a version of stratified comprehension which is restricted to those instances in which no type is assigned to a variable which is higher than the type which would be assigned to the set being constructed. This corresponds to a restriction on the impredicative formation of sets in *TT*. If the additional restriction is imposed that variables of the same type as the set being constructed must be parameters (must not be bound), we obtain the theory *NFP* (predicative *NF*). The versions of *TT* which correspond to these we will call *TTI* and *TTP*, respectively. The fragments of *NF* are equivalent to ambiguous or "tangled" versions of the corresponding type theories.

NFI and *NFP* are weak theories. The consistency of *NFI* is a theorem of third-order arithmetic, while the consistency of *NFP* is a theorem of first-order arithmetic. *NFI* interprets second-order arithmetic, while *NFP* interprets bounded arithmetic with exponentiation. *NFP* or *NFI* is strengthened to *NF* by the addition of the Axiom of Set Union; it is easy to see that for any stratified formula ϕ, we can construct $\mathcal{P}_1^n\{\{x \mid \phi\}\}$ in *NFP* (and so in *NFI*) for large enough n by choosing n so that the type of the set being constructed will be higher than the type of any bound variable in ϕ. Iterated applications of set union will then recover the desired set. The finite axiomatization for *NFU* + Infinity (axioms 1 through 16) can be adapted to *NFP* by dropping Set Union and strengthening extensionality; a type-level ordered pair can be defined in *NFP* as in *NF*.

NFP and *NFI* prove Infinity; if Union is true, we are in *NF*, and Specker's proof works; if Union is not true, observe that all finite sets have unions, so there must be an infinite set.

NFI and *NFP* invite applications which depend on the fact that they are weak theories which are very expressive in practical terms. *NFP* in particular is equivalent to a fragment of arithmetic (which one precisely is not known), but has the practical expressiveness of a set theory; it appears to be possible to do a good deal of elementary mathematics in this system. The author has shown the consistency of very powerful extensions of *NFI* in [16].

7 Comparison of Zermelo-Style Theories and Quine-Style Theories

We consider the possibility of application of the theories with stratified comprehension. NF_3 is too hard to work with to be a useful set theory, although it has an interesting role in the model theory of TT. *NFI* and *NFP* are also not practical set theories for general use, although they may have some interest precisely because they are weak systems. *NFP* may have an interesting relationship to systems of weak arithmetic currently being studied. It is also interesting to note that *NFP* is essentially Russell's original theory of types without the Axiom of Reducibility (basically equivalent to the Axiom of Set Union in modern formulations) and with typical ambiguity construed as identity; it is curious that collapsing the type structure enables us to dispense with the Infinity Axiom. Strong extensions of *NFI* would be hard to work in due to the failure of the Axiom of Set Union.

NF itself is strong enough to be a useful set theory, although the failure of the Axiom of Choice and the inability (so far) to produce interesting relative consistency and independence results makes it inconvenient. We do not believe that it is a reasonable "working set theory"; we have unmasked the special assumptions which distinguish it from *NFU* and force the failure of Choice in our discussion of "tangled type theory," and they seem to be of merely technical interest, and not worth the inconvenience of the failure of Choice, not to mention the opaque model theory! Our attitude might be different if *NF* had strong consequences of mathematical interest, but this seems not to be the case; there is no reason to believe that *NF* is any stronger than TT + Infinity or *NFU* + Infinity. If it were shown that *NF* is not consistent, this would not disturb us at all (although we would find it surprising): it would imply that the existence of urelements was provable in *NFU* without Choice, that the Axiom of Set Union could be disproven in *NFP*, and that there could be no type-level ordered pair in NF_3; it would also mean that the various theories with stratified comprehension would end up seeming more disconnected from each other than now seems likely.

The prime candidate for a useful working set theory of this type is *NFU* + Infinity + Choice, or the theory of Axioms 1 through 16 and 18 above. To this it immediately proves convenient to add the Axiom of Counting, giving the system of Axioms 1 through 18 above, the system of Rosser's *Logic for Mathematicians* modified to allow urelements, with the ordered pair as a primitive, and with the full Axiom of Choice. Axiom scheme 19 or other axioms could be used to get additional strength; the investigation of natural strong axioms of infinity for *NFU* is a possible area of research. This is a practical set theory, as we hope we have indicated above. Its obvious advantages are the ability to recover most of the constructions of naive set theory, including the natural definitions of cardinal and ordinal number, while retaining the ability to do the constructions of a Zermelo-style theory in ways we have indicated. The necessity of attention to stratification, if it is a disadvantage, is a slight one; much mathematical work is actually stratified in any case, and the author's experience indicates that it is not very hard to learn to keep track of relative type all the time, and it is sometimes instructive to do so. It is also worth noting that the constructions of the theory of sets and functions which are easy to implement in category theory are the stratified ones; elements are frequently represented by their constant functions in category-theoretic treatments for reasons similar to the reasons we introduce type-raising operations into constructions in *NFU*. A more serious difficulty is the odd model theory of *NFU*; there are externally countable proper classes and the ordinals in a model are not well-ordered. This requires getting used to, but is not unprecedented. The "internal set theory" approach to nonstandard analysis also involves small proper classes, for example (see [26]); in our motivation of the Axiom of Small Ordinals, there is some interaction with nonstandard analysis in the most general sense. It should be observed that membership is not a relation in *NFU*, so a *model* of *NFU*, in which the primitive predicates must be represented by sets, can be expected to behave oddly.

There is another problem for the Quine-Jensen theory which is more profound. This is the issue of philosophical motivation. The set theory of Cantor and Zermelo, and its extensions, are founded on the metaphor of the iterative construction of the universe of set theory in stages indexed by the ordinals. This metaphor helps us to convince ourselves of the reliability of these theories (see [20]). The historical motivation of the Quine-Jensen theories is a syntactical trick, and this fact helped to contribute to the error made at the outset of choosing strong extensionality, which has made the theories of this kind seem excessively difficult and mysterious (because *NF* is indeed difficult and mysterious!). Indeed, Quine seems himself to admit that this was an error in his remarks accompanying [19].

Jensen's proof and the variation using automorphisms (originally due to Boffa in [2], although implicit in [19]) give us reason to believe that *NFU* is reliable and has powerful extensions paralleling the extensions of the usual set theory with equal reliability. But this gives no independent interest to *NFU*. However, the fact that *NFU* recovers the details of its own model construction internally without recourse to the model theoretic machinery needed in the usual set theory for the same purpose should cause one to recover some confidence in the autonomy of the approach. It is interesting to observe that it seems to be technically easier to interpret set theory of the usual type in extensions of *NFU* than vice versa; one may be gaining practical expressive power by working in *NFU*.

We attempt a philosophical motivation of stratified comprehension. Consider the notion of *set* as being analogous to an abstract data type in computer science. The idea behind abstract data types is that the user of a concrete representation of a class of abstract structures should have access only to those properties of the concrete representation which represent properties of the abstract objects themselves, as specified by the designer of the abstract type. Suppose that we have a universe of objects, and we wish to implement classes of objects (pure extensions, not yet identified with any objects) as objects themselves. The abstract structures being represented are very simple: they are pure extensions. But the implementation of each class as a set has an additional feature: it is identified with an object. The implementation of each class is not a pure extension (collection of objects), but an extension with an object label attached. The analogy to abstract data types dictates that any feature of an implemented set we use should be a feature of the pure extension, not of the structure "label + extension" which implements it.

Now consider the "property" of sets "x is an element of itself". This property, unlike its notorious negation, is not paradoxical. But we assert that it is illegitimate by the criterion inidcated above: it is a property of the structure "label + extension" for each set, not a property of the extension considered by itself. One way to see this is to consider that a permutation technique could be used to reshuffle the "labels" attached to classes without changing which classes are implemented, and an extension previously labelled by one of its elements could come to be labelled by an object which is not one of its elements. Thus we can reject this property (and its paradoxical complement) on *a priori* grounds before ever considering the paradoxes.

The difficulty with "$x \in x$" as a property of x is that it involves x in two different roles which are linked in a way which depends on arbitrary details of the implementation of classes as labelled extensions rather than on the intrinsic properties of objects or classes considered in themselves. Analysis reveals that there are further roles beyond the roles of "object associated with an extension (set)" and "bare object (element of a set)" revealed in this example. An object can be associated with a collection of extensions (set of sets) via the association of each of the elements of its associated extension with an extension; similarly, an object can be associated with a set of sets of sets, and so on through the usual type hierarchy. Any two of these roles for an object are linked only by the details of the labelling used to implement sets, so an object should be mentioned in the definition of an intrinsic property of sets in no more than one of these roles. But this is exactly the criterion of stratification!

The philosophical motivation is approached from another angle in the discussion of an abstract model of computation in the next section.

The question of extensionality now enters. Strong extensionality amounts to the assertion that every object is a set, which seems far from natural; *NFU* seems like the more reasonable position to start from. It is interesting to observe that Marcel Crabbé has shown (in [5]) that the theory *SC* whose sole axiom scheme is stratified comprehension interprets *NFU*, so the weak extensionality of *NFU* adds no consistency strength to the theory, while strong extensionality adds at least Infinity. There are other philosophical questions about the impredicativity of the full stratified comprehension scheme, which are equally a problem for *TT* and the usual set theory, and best answered by observing that *NFU* is demonstrably no more dangerous than these theories.

A nice observation, in view of the illustrative role of permutation techniques in the argument above, is that Forster has shown in [9] that the notion that a formula is invariant under all set permutations of the definition of membership is expressible in *NF(U)* (for each formula individually) and can be used as the "necessity" operator of a modal logic satisfying nice axioms. The notion that a formula is true under some permutation serves as the "possibility" operator. All stratified sentences are invariant under permutations, as are some unstratified sentences, such as the Axiom of Counting. If a slightly more general class of permutations, including some external permutations (the "set-like" permutations defined in [9]) is allowed, the invariant sentences become precisely the stratified sentences.

The basic question for any set theory is "When is there the collection of all objects with a given property?". The theory of Cantor and Zermelo is oriented toward criteria which ensure that it will be safe to *collect* objects (limitation of size, well-foundedness), while the theory of Quine and Jensen takes the riskier course of defining criteria for sensible *properties*. The Quine-Jensen approach allows larger collections, and it is possible to arrange for it to allow all the collections allowed by the Cantor-Zermelo approach as well.

It should be observed that the fact that *NFU* has a universe, a largest cardinal and a set of all ordinals does not mean that the Absolute Infinite inconsistent totalities (to use Cantor's terminology) are quite invisible from this standpoint. The "limit" of the strongly Cantorian ordinals in *NFU* is closely related to the Absolute Infinite "limit" of all the ordinals in the usual set theory; the permutation which interchanges $\{\beta \mid \beta < \alpha\}$ with $\mathrm{T}\{\alpha\}$ for each ordinal α can be used to redefine membership in such a way that exactly the strongly Cantorian ordinals become von Neumann ordinals (how many von Neumann ordinals one gets is a function of how strong an extension of *NFU* one is working in). The fact that the (consistent) limit of *all* ordinals is "above" the analogue of the Absolute Infinite is a little surprising, but can be motivated by the idea used to motivate the Axiom of Small Ordinals: the ordinals of *NFU* are an initial segment of a nonstandard extension of the standard ordinals.

The most serious challenge to the legitimacy of *NFU* and related theories as set theories has to do with the existence of the external descending sequence in the ordinals and analogous phenomena; it could be claimed that this signals something essentially wrong with the definition of well-ordering in *NFU*. But this can only be construed as an objection if one assumes at the outset the doctrine of "limitation of size" as the criterion for sethood, thus implicitly choosing Zermelo-style set theory. A relation is a well-ordering if each subset of its domain has a least element, in *NFU* as in *ZFC*; the range of the external descending sequence of ordinals does not meet the criteria for sethood in *NFU*. If we require of a well-ordering that each *subclass* of its domain have a least element, we have (under the assumption of the Axiom of Small Ordinals and treating classes as being defined by formulae) restricted ourselves to the proper class of Cantorian ordinals, which are indeed more closely analogous to the proper class of ordinals of the usual set theory. There is no reason to expect a class defined using quantification over all classes to be a set; if one protests that the subclasses of a set must be a set, one is begging the question: "each subclass of a set is a set" is a formulation of the axiom of comprehension of Zermelo set theory, which is admittedly incompatible with *NFU* (or any set theory with a universal set or set of all ordinals)!

8 Stratified Theories of Functions and Potential Applications in Theoretical Computer Science

The "philosophical motivation" for *NFU* can be recast more explicitly in terms of theoretical computer science (on a *very* abstract level) in the context of a theory of functions rather than of sets.

Let an abstract "machine" be understood as a set X of "addresses". Let the same set X be used as input or output "data". Then a "program" can be understood as being a function from X to X. Let a "state" of the machine be a function from the set of "addresses" to the set of "programs"; i.e., a function from X to X^X; the "state" function tells us which "programs" are stored in which "addresses". If Σ is the state of the abstract machine and x and y are "addresses," we define the "application" $x(y)$ of x to y as $\Sigma(x)(y)$, the result of supplying y as input to the program stored in address x.

It is convenient to have a pair on X: if x and y are addresses, (x, y) is an address. It is possible to pair programs as well: if P and Q are programs, $(P, Q)(x) = (P(x), Q(x))$ defines the pair (P, Q) as a program. It is convenient (and provably possible) to assume that $\Sigma(x, y) = (\Sigma(x), \Sigma(y))$; the pairing operations on addresses and on programs are parallel. This gives us the equation $(x, y)(z) = (x(z), y(z))$ for x, y, and z addresses as well.

Since we have pairing, we would like to have the projection functions stored in addresses π_1 and π_2, giving equations $\pi_i(x_1, x_2) = x_i$ for $i = 1, 2$. We would also like to encode equality of addresses: $\text{Eq}(x, y) = \pi_1$ if $x = y$ and π_2 otherwise, for some address Eq.

Finally, we would like to have "abstraction": if T is a term written using application, pairing, and (recursively) abstraction, plus any other constructions we might include in our language, we want an address $(\lambda x)(T)$ such that $\Sigma((\lambda x)(T))$ is the function which takes each a to $T[a/x]$ (the result of substituting a for x) for each a in X. If terms S and T have the same values for each value of the variable x, we postulate that $(\lambda x)(S)$ and $(\lambda x)(T)$ will be equal. It is also convenient and proves harmless to assume that $(\lambda x)(S,T) = ((\lambda x)(S), (\lambda x)(T))$.

It turns out that the system described so far is inconsistent, a result of Curry (see [6]). Let $\text{Neg} = (\lambda x)(\text{Eq}(x, \pi_1)(\pi_2, \pi_1))$. Neg sends π_1 to π_2 and each other x to π_1; clearly, it has no fixed point. Now consider $R = (\lambda x)(\text{Neg}(x(x)))$. $R(R) = \text{Neg}(R(R))$ by the definition of abstraction, which is impossible. The R is in honor of Russell, for this "Curry paradox" is another version of Russell's paradox.

Reflections similar to those above on the motivation of stratified comprehension give an indication why we should reject the definition of R on *a priori* grounds, before we even consider the possibility of paradox. Since paradoxical functions make us queasy, we will use $\Delta = (\lambda x)(x(x))$ as our representative "bad" function. Δ tells us to take our input (an address) and apply the program stored in the address to the address itself. But there is a violation of the security of the data type "program" here: it should be even clearer in this context that an address x is not an intrinsic feature of the program $\Sigma(x)$ stored there, nor is the program stored in x an intrinsic feature of x as an address. This association is an essentially arbitrary feature of the state Σ of our abstract machine, not a feature of the "program" considered in itself. Thus, Δ is seen not to be a legitimate operation on programs.

Note that we can represent not only bare addresses and programs, but also functions from programs to programs (second-order programs), functions from second-order programs to second-order programs (third-order programs) and so on. In general, we can define a 0th order program as an address and an $(n+1)$-st order program as a function from nth order programs to nth order programs. Objects of each of these types are represented on the abstract machine, and the relationship between the mth order program and the nth order program for $m \neq n$ stored in the same address x is not a feature of either abstract object in itself, but in relation to the essentially arbitrary state of the machine, so should not be reflected in the specifications of legitimate operations on these objects. (Objects of types like "function from first-order programs to second-order programs" can be encoded in this scheme, but we do not discuss this here.)

The criterion for a legitimate operation (of any order) is that it be possible to assign a fixed order n indicating a role as an nth order program to each term in the definition of the operation in a sensible way. Δ fails this test: in $x(x)$, if we assign the first occurrence of x the order n, we want to assign the second occurrence the order $n-1$.

If we assign a term $x(y)$ order n, it should be clear that we want to assign y the same order n and x the order $n+1$. If we assign a term (x, y) order n, we assign the same order to x and y. If we assign order n to an abstraction $(\lambda x)(T)$, we assign type $n-1$ to T, and we expect to assign no type other than $n-1$ to x inside T; if this expectation is disappointed, the abstract does not represent a legitimate operation. These rules allow us to assign orders to all subterms of a term in our language once we give the term itself a sufficiently high order; they also allow us to recognize terms that need to be rejected as ill-formed. The restrictions on well-formedness of abstraction terms correspond to the stratification restrictions in *NFU*. As in *NFU*, the objects we are working with are all actually of the same sort; they are all "addresses". It is useful to observe that the relation between addresses of representing the same nth order object is definable for each n.

We have shown in [13] that this system of "stratified λ-calculus" interprets and is interpreted by *NFU* + Infinity. It should be noted that this degree of strength depends on a weak extensionality principle supported by the possibility of substitution of equals for equals in abstracts: if $S = T$ for all values of x, then $(\lambda x)(S) = (\lambda x)(T)$. If full extensionality (if $f(x) = g(x)$ for all x, then $f = g$) is added, we obtain a theory equiconsistent with *NF*. Models of the theory with weak extensionality can be constructed in a way analogous to the construction of models of *NFU* above, mod technical considerations (it is necessary to arrange for X^X to "include" X). There is also a subset of extensional stratified λ-calculus which is analogous to and equiconsistent with *NFI*; the predicative stratified λ-calculus analogous to *NFP* apparently cannot be shown to be equiconsistent with *NFP* by the same methods but remains interesting.

These stratified λ-calculi can also be presented as systems of synthetic combinatory logic. We present the synthetic combinatory *TRC* (for "type-respecting combinators") shown to be equivalent to *NF* in our [13], along with some of its fragments. An advantage of these systems, analogous to the advantage of our axioms 1 through 16 for *NFU* + Infinity, is that the axioms and term formation rules make no mention of stratification.

TRC is a first-order theory with the following term formation rules: π_1, π_2, Abst, and Eq, and variables are atomic terms; if f and g are terms, $f(g)$, (f, g), and $\mathrm{K}[f]$ are terms; write $f(g, h)$ for $f((g, h))$ and $\mathrm{K}^n[f]$ to abbreviate n-fold application of the K constructor. The axioms of *TRC* are as follows:

Axiom 1. $\pi_i(x_1, x_2) = x_i$ (for $i = 1$ or 2)

Axiom 2. $(\pi_1(x), \pi_2(x)) = x$

Projection operators are projection operators and pairing is surjective (the latter is also a convenient assumption to add to the system above).

Axiom 3. $\mathrm{K}[f](g) = f$

Axiom 4. $(f, g)(h) = (f(h), g(h))$

Axiom 5. $\mathrm{Abst}(f)(g)(h) = f(\mathrm{K}[h])(g(h))$

Axioms from which a definition of abstraction can be recovered. K is the constant function constructor (note that it is *not* a function). Products of functions are pairs of functions. Abst is a stratified analogue of the **S** combinator in the usual untyped, unstratified combinatory logic. An identity function Id can be defined as (π_1, π_2), and axioms 2 and 4 used to verify that $\mathrm{Id}(x) = x$.

Axiom 6. $\mathrm{Eq}(f, g) = \pi_1$ if $f = g$, π_2 otherwise.

Axiom 7. If $f(x) = g(x)$ for all x, then $f = g$.

Axiom 8. $\pi_1 \neq \pi_2$

This completes the axioms of *TRC*. If axiom 7 of extensionality is dropped, it needs to be replaced by a weak extensionality axiom involving a new atom Ext to retain the strength of *NFU* + Infinity or the system above, along with some additional axioms for Ext. Ext sends each function to a canonical function with the same extension. These axioms follow:

Axiom $7'$**.** $f(x) = g(x)$ for all x iff $\mathrm{Ext}(f) = \mathrm{Ext}(g)$.

Axiom E. $\mathrm{Ext}(\mathrm{Ext}(f)) = \mathrm{Ext}(f)$; $\mathrm{Ext}(f, g) = (\mathrm{Ext}(f),\mathrm{Ext}(g))$

The system *TRCL* which results if these axioms replace axiom 7 is equivalent to the stratified λ-calculus without extensionality or to *NFU* + Infinity.

The axioms and primitive notions of *TRC* or *TRCL* are quite natural, with the exception of the object Abst and its axiom. If we drop the axiom Abst and its axiom (retaining axiom 7) and replace it with a new atom Comp with axiom

Axiom $5'$**.** $\mathrm{Comp}(f, g)(h) = f(g(h))$

we obtain a synthetic combinatory logic *TRCP* equivalent to the predicative version of the stratified λ-calculus with full extensionality. The replacement of the peculiar Abst with the natural notion of composition on functions makes *TRCP* very appealing; like *NFP*, it is weaker than arithmetic, but it is strong enough to do a lot of elementary mathematics. A version *TRCI* which is precisely equivalent to *NFI* in consistency strength and expressive power is obtained by adding an axiom scheme of a quite forbidding appearance to *TRCP*; equivalence results relating *TRCP* and *NFP* are blocked by the failure of the Schröder-Bernstein theorem in *TRCP* and *NFP*. For details on the systems *TRCP* and *TRCI*, see our [15].

In the synthetic theories, it is possible to give an inductive definition of $(\lambda x)(T)$ and prove the stratified abstraction scheme "$(\lambda x)(T)(x) = T$" for each T in which x occurs with no relative type other than that of T for a suitable definition of relative type (and with other restrictions in *TRCI* and *TRCP*, restricting impredicativity). One can then use characteristic functions to interpret sets and interpret set theories with stratified comprehension. The interpretation of these systems of combinatory logic and λ-calculus in the relevant fragments of *NF* is made technically involved by the axiom $(f, g)(h) = (f(h), g(h))$ which identifies products and pairs of functions, but the gain in convenience is considerable (we would otherwise need an atomic term denoting the product operation).

We think that these systems of combinatory logic and λ-calculus may find applications in computer science, where closely related typed λ-calculi are already in use. These systems involve a limited kind of polymorphism.

We present observations relating the notion "strongly Cantorian" from *NFU* to the notion of a "data type" from computer science, via our abstract model of computation. Since the model is extremely abstract, the correspondence is necessarily rather vague, but it may be suggestive. It is especially interesting because the concept "strongly Cantorian" corresponds to nothing in the usual set theory.

Recall that a set A is strongly Cantorian exactly when the map $(\iota|A)$ which sends elements of A to their singletons exists. "Big" sets like the universe are not strongly Cantorian. *NFU* + Infinity by itself does not prove the existence of any strongly Cantorian sets except the sets of standard finite cardinality. The Axiom of Counting causes all finite sets, $\mathcal{N}$, and many larger sets to become provably strongly Cantorian. The Axiom of Small Ordinals has an even more powerful effect.

Since we are now working in a theory whose primitive notion is function application rather than set membership, we need to indicate what we mean by a "set". We regard characteristic functions (functions f such that $f(x) = \pi_1$ or $f(x) = \pi_2$ for all x) as the representatives of sets in these systems. We define "$x \in A$" as "A is a characteristic function and $A(x) = \pi_1$". Note that the type differential between x and A is the same as it would be in *NFU*. It is straightforward to show that stratified λ-calculus or *TRCL* interprets *NFU* + Infinity with this definition of membership (objects which are not characteristic functions are interpreted as urelements).

Instead of defining "strongly Cantorian" in terms of the singleton map, we define it in terms of the (equally non-existent) constant function operator. We say that a set A is strongly Cantorian if the function $(\mathrm{K}|A)$ which sends each $a \in A$ to $\mathrm{K}[a]$ and each $x \notin A$ to π_2 (a default value) exists, and so does its "inverse" $(\mathrm{K}|A^{-1})$, which sends $\mathrm{K}[a]$ to a for $a \in A$ and other objects to the default value. There are technicalities involved in making this work correctly: an axiom should be added which asserts that each "set" of ordered pairs which would be interpreted as a function in a set theory corresponds to an actual function; such an axiom does not strengthen these theories essentially. It is easy to show that this definition of "strongly Cantorian" is equivalent to the usual one.

The class of strongly Cantorian sets is closed under operations like disjoint union, Cartesian product, the formation of function spaces, and other constructions on types used in computer science, and contains sets of standard finite size. If the Axiom of Counting holds, it contains $\mathcal{N}$ and is closed under various inductive constructions. It is not absurd to think that this corresponds to a notion of "type".

As we observed above in the set theory context, variables restricted to strongly Cantorian sets can evade stratification restrictions. The reason for this is that an instance of x can have its type raised by substituting $(\mathrm{K}|A)(x)(\mathrm{Id})$ for it or lowered by substituting $(\mathrm{K}|A^{-1})(\mathrm{K}[x])$ for it; either of these is equal to x as long as x is in A. We can raise or lower the types of each instance of a variable x restricted to A until all have the desired type and abstraction relative to x can be carried out.

For example, we can define $(R|A)$ for each strongly Cantorian set A, where R is the paradoxical operation which sends x to $\mathrm{Neg}(x(x))$. $\mathrm{Neg}(x(x)) = \mathrm{Neg}((\mathrm{K}|A^{-1})(\mathrm{K}[x])(x))$, and $(\lambda x)(\mathrm{Neg}((\mathrm{K}|A^{-1})(\mathrm{K}[x])(x))$ is well-formed (all x's have the correct relative type).

The interpretation of this phenomenon in the light of our abstract model of computation motivates the identification of the notion of "strongly Cantorian set" with the notion of "data type". The reason that the stratification restrictions hold in general is that it is illegitimate to refer to the relation between programs and the addresses in which they are stored. The map $(\mathrm{K}|A)$ can be regarded as coding the correlation between elements of A as addresses and elements of A as programs in a way which our model can accept: $(\mathrm{K}|A)$ associates $\mathrm{K}[x]$ (a program which always returns x, understanding x as an address) with x (which must here be understood as a program). $(\mathrm{K}|A)$ encodes the way in which the abstract objects interpreted by elements of A are stored in addresses in the "memory" of our "machine"; a class of objects with which we have associated a protocol for storing them in memory is a good description of what a "data type" is, on a high level of abstraction. We combine this analogy and the empirical observation that the class of strongly Cantorian sets is closed under the usual type constructors to support our basic claim.

It is amusing (and perhaps interesting as well) to observe that, if we accept this analogy, the two ways in which the notion of "type" are reflected in *NFU* are orthogonal to one another: there is the notion of relative type used in stratification and the notion of a strongly Cantorian set, understood as "data type". A strongly Cantorian set is a domain on which the restrictions imposed by relative typing can safely be ignored!

We believe that *TRCL*, a system with the strength of Russell's theory of types which is an untyped system of synthetic combinatory logic, a theory of functions of universal domain, may be regarded as fulfilling Curry's program for the foundation of mathematics, if our philosophical argument for stratified functional abstraction is accepted as a critique of his original requirement of unrestricted functional abstraction (see the introduction to [6]).

9 Back to the Sources

Since interest in this field has hitherto been rather limited, the number of easily accessible sources is not great. Quine himself wrote the paper [24] introducing the whole idea and the book [25], which is based on an extension of *NF* with an impredicative theory of proper classes. Neither of these sources is entirely reliable. In [24], Quine gave an invalid argument for the Axiom of Infinity. In [25], Quine claimed that he had overcome the problems of *NF* with mathematical induction on unstratified conditions by defining $\mathcal{N}$ as the intersection of all inductive *classes* instead of the intersection of all inductive sets. This does indeed enable mathematical induction on unstratified conditions, but it also means that one can no longer establish that $\mathcal{N}$ is a set rather than a proper class. This is embarrassing in [25], where Quine later constructs the reals in a way which requires certain sets of rational numbers to be elements. It can be repaired by adding an axiom which asserts that $\mathcal{N}$ thus defined is indeed a set, which amounts to adding strong mathematical induction to *NF* (a little more than this because the class theory in [25] is impredicative). Also, the theory in the original version of [25] was inconsistent (a fact which has led to persistent rumors that *NF* itself is inconsistent), because it allowed sets to be defined using quantification over classes; this was corrected by Wang and the correction has appeared in subsequent editions.

A better source for studying *NF* is [27], Rosser's *Logic for Mathematicians*, especially in conjunction with the dicta we have given above for converting his system to *NFU*. Make the type level pair a primitive notion, weaken extensionality to allow urelements, and assume the full form of the Axiom of Choice (Rosser discusses the relationships between various forms of the Axiom of Choice, but only finally commits himself to the denumerable Axiom of Choice; Specker's subsequent results did not invalidate this system). This is a good book with a public relations problem which attention to results about *NFU* can solve. One difficulty with this book is that its notation is nonstandard for the usual notions of set theory; but a familiarity with Rosser's notation is useful for reading some of the papers on *NF*.

The watershed papers about *NF* and related systems from our standpoint are [29], in which Dana Scott first applied the method of permutations to give relative consistency and independence results for *NF* (but only for unstratified sentences) (this work was extended by C. W. Henson), Specker's papers [30], in which the Axiom of Choice was disproved and [31] in which the basic model theory results were given, and the three papers [19], [10], and [4], in which *NFU*, NF_3, and *NFI*, respectively, were shown to be consistent.

Until very recently the only book-length treatment of *NF* postdating the results of Specker was Forster's [8], which was not readily available and did not discuss the consistent subsystems. Forster has remedied this situation with his excellent book *Set theory with a universal set* ([9]), which, in spite of the generality of its title, is mostly about *NF*. It does describe the related systems

NFU and NF_3 (not *NFI* or *NFP*) and give the basic results. It also has an excellent bibliography, to which we refer the reader. We hope to write another book, in collaboration with other workers, which will give a detailed treatment of the consistent fragments of *NF* and the related λ-calculi.

References

[1] Peter Aczel, *Non-well-founded sets*, CSLI lecture notes, Stanford University (distributed by Chicago University Press), 1988.

[2] Maurice Boffa, "Entre *NF* et *NFU*," *Comptes Rendues*, series A, 277 (1973), pp. 821–2.

[3] J. H. Conway, *On Numbers and Games*, Academic Press, London, 1976.

[4] Marcel Crabbé, "On the consistency of an impredicative subsystem of Quine's *NF*," *Journal of Symbolic Logic*, 47 (1982), pp. 131–36.

[5] Marcel Crabbé, "On *NFU*," Preprint.

[6] H. B. Curry and R. Feys, *Combinatory Logic*, Vol. I, North Holland, Amsterdam, 1958.

[7] Daniel Dziergowski, "Le théorème d'ambiguïté et son extension à la logique intuitionniste" (in French), doctoral thesis, Catholic University of Louvain, 1993.

[8] Thomas Forster, "Quine's *NF*, an introduction," *Cahiers du Centre de Logique*, no. 5, Louvain-la-Neuve, 1983.

[9] T. E. Forster, *Set theory with a universal set, an exploration of an untyped universe*, Oxford logic guides, no. 20, 1992.

[10] V. N. Grishin, "Consistency of a fragment of Quine's *NF* system," *Sov. Math. Dokl.*, 10 (1969), pp. 1387–90.

[11] T. Hailperin, "A set of axioms for logic". *Journal of Symbolic Logic*, 9 (1944), pp. 1–19.

[12] Leon Henkin, "The representation theorem for cylindrical algebras", in Thoralf Skolem *et. al.*, *Mathematical Interpretations of Formal Systems*, North-Holland, Amsterdam, 1955.

[13] M. Randall Holmes, "Systems of Combinatory Logic Related to Quine's 'New Foundations'," *Annals of Pure and Applied Logic*, 53 (1991), pp. 103–33.

[14] M. Randall Holmes, "The Axiom of Anti-Foundation in Jensen's 'New Foundations with Ur-Elements'," *Bulletin de la Societe Mathematique de la Belgique*, series B, 43 (1991), no. 2, pp. 167–79.

[15] M. Randall Holmes, "Systems of Combinatory Logic Related to Predicative and 'Mildly Impredicative' Fragments of Quine's *NF*," *Annals of Pure and Applied Logic*, 59 (1993), pp. 45–53.

[16] M. Randall Holmes, "The equivalence of *NF*-style set theories with "tangled" type theories; the construction of ω-models of predicative *NF* (and more)," Preprint.

[17] M. Randall Holmes, "Strong axioms of infinity in *NFU*," in preparation. Draft is available on request.

[18] C. W. Henson, "Permutation methods applied to *NF*," *Journal of Symbolic Logic*, 38 (1973), pp. 69–76.

[19] Ronald Bjorn Jensen, "On the consistency of a slight (?) modification of Quine's 'New Foundations'," *Synthese*, 19 (1969), pp. 250–63.

[20] Penelope Maddy, "Believing the Axioms, I," *JSL*, 53 (1988), pp. 481-511, and "Believing the Axioms, II," *JSL*, 53 (1988), pp. 736-64.

[21] Colin McLarty, "Failure of Cartesian closedness in *NF*," *Journal of Symbolic Logic*, 57 (1992), pp. 555–6.

[22] S. Orey, "New Foundations and the axiom of counting," *Duke Mathematical Journal*, 31 (1964), pp. 655–60.

[23] J. F. Pabion, "TT_3I est équivalent à l'arithmétique du second ordre" (in French), *Comptes Rendues*, series A, 290 (1980), pp. 1117–8.

[24] W. V. O. Quine, "New Foundations for Mathematical Logic," *American Mathematical Monthly*, 44 (1937), pp. 70–80.

[25] W. V. O. Quine, *Mathematical Logic*, second edition, Harvard, 1951.

[26] Alain Robert, *Nonstandard Analysis*, John Wiley and Sons, 1988.

[27] J. Barkley Rosser, *Logic for Mathematicians*, second edition, Chelsea, New York, 1973.

[28] Alfred North Whitehead and Bertrand Russell, *Principia Mathematica (to *56)*, Cambridge University Press, 1967.

[29] Dana Scott, "Quine's individuals," in *Logic, methodology and philosophy of science*, ed. E. Nagel, Stanford, 1962, pp. 111–5.

[30] E. P. Specker, "The axiom of choice in Quine's 'New Foundations for Mathematical Logic'," *Proceedings of the National Academy of Sciences of the U. S. A.*, 39 (1953), pp. 972–5.

[31] E. P. Specker, "Typical ambiguity," in *Logic, methodology, and philosophy of science*, ed. E. Nagel. Stanford, 1962.