Notre Dame Journal of Formal Logic
Volume 40, Number 2, Spring 1999

Subsystems of Quine's "New Foundations" with Predicativity Restrictions

M. RANDALL HOLMES

Abstract This paper presents an exposition of subsystems NFP and NFI of Quine's NF, originally defined and shown to be consistent by Crabbé, along with related systems TTP and TTI of type theory. A proof that TTP (and so NFP) interpret the ramified theory of types is presented (this is a simplified exposition of a result of Crabbé). The new result that the consistency strength of NFI is the same as that of PA_2 is demonstrated. It will also be shown that NFI cannot be finitely axiomatized (as can NF and NFP).

1 Introduction This paper has two aims, both related to subsystems of Quine's "New Foundations" (hereinafter NF) proposed and shown to be consistent by Crabbé in [1]. The first aim is largely expository. In [2], Crabbé showed that the predicative version of the simple theory of types (Russell's theory of types as simplified by Ramsey) is equiconsistent with (a formalization of) the ramified theory of types. However, Crabbé's presentation is quite complex and hard to follow. We give a much more direct demonstration of this equivalence (also owing much to Crabbé, we must hasten to add). It follows from this that predicative NF is also equiconsistent with the ramified theory of types (with the axiom of infinity); this should be of interest because predicative NF is formally a much simpler theory.

The second aim is to present a new research result. In [1], Crabbé demonstrated the consistency not only of predicative NF (NFP) but also of an impredicative fragment NFI of NF. Crabbé showed that Peano Arithmetic proves the consistency of predicative NF, and that third-order arithmetic, PA_3, proves the consistency of NFI. We prove the sharper result that the consistency strength of NFI is exactly that of second-order arithmetic, PA_2.

***2 Simple type theory and* NF** The simple theory of types TT is a many-sorted first-order theory with membership and equality. The sorts, called *types*, are indexed

Received December 23, 1999; revised April 10, 2000

by the natural numbers. Informally, type 0 is inhabited by "individuals" of an unspecified nature; objects of type $n+1$ are sets of type n objects. This informal intuition is embodied in the type conventions for atomic formulas: $x = y$ is well-formed if and only if the variables x and y are of the same type, while $x \in y$ is well-formed if and only if the type of y is the successor of the type of x. (We stipulate that each variable in the language of TT has a type, and that we have a countable supply of variables of each type, but we avoid the notational ugliness of explicit type superscripts on each variable.)

The axioms of TT include an extensionality scheme:

axiom of extensionality $(\forall xy.(\forall z.z \in x \equiv z \in y) \equiv x = y)$

This is a scheme because different versions are needed with variables z of each type (the types of x and y will be one higher).

The axioms are completed by the scheme of comprehension:

axiom of comprehension $(\exists A.(\forall x.x \in A \equiv \varphi))$, for each formula φ in which the variable A does not occur free.

An informal way to state the instance of comprehension corresponding to a formula φ (and variable x) is "$\{x \mid \varphi\}$ exists".

One frequently adjoins an axiom of infinity to TT and sometimes the axiom of choice. TT with infinity and choice is more than adequate for all classical mathematics outside of set theory.

An important phenomenon in TT is "typical ambiguity". Any axiom of TT, and so any theorem, remains an axiom or theorem if the type of each variable appearing in it is raised by one (and so by any uniform amount). Any object which can be defined in TT has precise analogues in each higher type. For example, one can define natural numbers using Frege's approach: 3, for example, is the set of all sets with three elements (the notion of having three elements can be defined in first-order logic; there is no circularity). But one obtains numerals 3 in each type above 2: the set of all sets of three type 0 objects is a type 2 object, while the set of all sets of three type 1 objects is a type 3 object, and so forth. We should be careful not to say that we have a sequence of different numerals 3: these are objects of different sorts and so cannot be compared pairwise or collected into a sequence!

The phenomenon of typical ambiguity can (apparently) be exploited to simplify the theory. The first-order theory with equality and membership obtained by dropping all indications of type from the axioms of TT (in such a way as to create no identifications of variables originally of different types), considered by Quine in [7], and called "New Foundations" or NF, is, as far as anyone knows, consistent (its consistency relative to generally accepted systems of set theory remains an open question).

The extensionality scheme of TT corresponds to a single extensionality axiom in NF. The comprehension scheme of TT (which we can informally describe as "$\{x \mid \varphi\}$ exists") does *not* translate to the naive comprehension scheme which asserts that "$\{x \mid \varphi\}$ exists" for each formula φ in the language of set theory. The point is that not all formulas φ can be obtained by suppressing distinctions of type in formulas of TT.

It is possible to describe the comprehension scheme of NF without referring to TT.

Definition 2.1 Let φ be a formula of the language of set theory and let σ be a function from variables (as syntactical items) to integers. We call σ a *stratification of* φ if and only if for each atomic subformula '$x = y$' of φ we have $\sigma(`x') = \sigma(`y')$ and for each atomic subformula '$x \in y$' of φ we have $\sigma(`x') + 1 = \sigma(`y')$. We call a formula φ *stratified* if there is a stratification of φ.

The comprehension scheme of NF (usually called the axiom (scheme) of stratified comprehension) asserts that "$\{x \mid \varphi\}$ exists" for each stratified formula φ. The definition of stratification can be extended to accommodate any additional predicates or functions one could adjoin to TT.

We briefly recapitulate the formal definition of the theory NF. NF is a first-order unsorted theory with primitive predicates of equality and membership. Its axioms are:

axiom of extensionality	$(\forall xy.(\forall z.z \in x \equiv z \in y) \equiv x = y)$
axiom of stratified comprehension	$(\exists A.(\forall x.x \in A \equiv \varphi))$, for each stratified formula φ in which the variable A does not occur free.

NF is an unusual set theory for two reasons. First of all, the universal set $V = \{x \mid x = x\}$ and other big sets (such as Frege natural numbers) exist in NF. Secondly, the axiom of choice can be disproved in NF (a result of Specker in [10]). Note that a paradoxical set like $\{x \mid x \notin x\}$ (the Russell class) is not provided by the stratified comprehension scheme because $x \notin x$ is not a stratified formula.

The presence of big objects, though unfamiliar to those used to Zermelo-style set theory, has been proved not to be problematic in any essential way. The variant NFU of NF proposed by Jensen in [6], in which extensionality is weakened to apply only to objects with elements (thus allowing many distinct objects without elements, which are called atoms or *urelements*—thus the U) has exactly the same comprehension scheme (and so the same big sets) and is known to be consistent relative to familiar set theories. Moreover, it is also consistent with infinity and choice.

The disproof of choice in NF cannot be replicated in any subsystem of NF which is known to be consistent. A corollary of the failure of choice is that the "axiom" of infinity is provable in NF; thus a lower bound on the consistency strength of NF (the best lower bound known) is that of TT + Infinity. Infinity is not provable in *NFU*, but it is provable in predicative NF, which is known to be consistent (and the proof of infinity in NFP makes essential use of the provability of Infinity in NF). The peculiarities of NF revealed by Specker's disproof of choice are profound, unlike the peculiarities of having big sets like the universe that are shared with NFU.

3 Predicative and 'mildly impredicative' theories The simple theory of types and NF are highly impredicative theories. That is, they allow the "definitions" of totalities by comprehension axioms to refer to totalities which include the set being "defined". For example, the comprehension axiom which asserts the existence of the set of natural numbers $\mathcal{N}$ describes $\mathcal{N}$ as the intersection of all sets of cardinal numbers which contain 0 and are closed under addition of 1. $\mathcal{N}$ itself is a set of cardinal numbers which contains 0 and is closed under addition of 1; it seems that we have "defined" $\mathcal{N}$ in terms of a class of sets to which it itself belongs.

For the record, we see no philosophical problem with the comprehension axiom which provides us with $\mathcal{N}$. We agree with Russell that if it were really to be regarded as a definition, it would be circular—the correct conclusion to draw is that comprehension axioms are not definitions!

Russell and Whitehead proposed that the paradoxes of set theory were due to impredicativity and proposed to recast their system of *Principia Mathematica* to avoid impredicativity. The original predicative type theory of Russell, called the ramified theory of types, is quite complex; we will present a formalization of this system below. A simpler predicative system is readily obtained by restricting the comprehension axioms of TT, and the first aim of our paper is to show that this system is in fact equivalent in a strong sense to the ramified theory of types; each system interprets the other readily.

We define a subtheory TTP of TT in order to satisfy scruples about predicativity. The extensionality scheme of TTP is the same as that of TT. The comprehension scheme is restricted in a way which ensures that no object can be "defined" (provided by a comprehension axiom) when the comprehension axiom involves quantification over the type to which the object belongs. An instance of the scheme "$\{x \mid \varphi\}$ exists" is excluded if φ includes a quantifier over the type to which $\{x \mid \varphi\}$ itself would belong (one higher than the type of x) or any higher type. Further, an instance of "$\{x \mid \varphi\}$ exists" is excluded if any object of type higher than that of $\{x \mid \varphi\}$ is mentioned. It is permitted for φ to refer to objects of the same type as $\{x \mid \varphi\}$; for example, the comprehension axiom "$\{x \mid x \in A \wedge x \in B\}$ exists" which "defines" $A \cap B$ is predicative and has parameters A and B at the same type as $A \cap B$.

To motivate these restrictions intuitively, imagine that the types are created in order. The definition of a type n object cannot involve quantifiers over type n or any higher type, because the creation of type n is not yet complete (so we do not know what might be true of *all* type n objects) and the creation of still higher types has not yet begun. The permission to use individual type n objects as parameters in the definition of other type n objects can be understood as reflecting the fact that objects are created in a certain order within the types; of course, no object of higher type can be used as a parameter because no objects of higher type have yet been created.

Thus, the formal restriction on the comprehension scheme of TTP is that a permitted comprehension axiom "$\{x \mid \varphi\}$ exists" is one in which the formula φ contains no bound variable with type higher than the type of x and no free variable (parameter) with type higher than the successor of the type of x.

The system of set theory obtained by suppressing the type system of TTP in the same way that NF is obtained from TT is called NFP or "predicative NF". Unlike NF this system is known to be consistent and in fact quite weak.

The comprehension scheme of NFP can be presented in terms of stratifications.

Definition 3.1 A stratification σ of a formula φ is said to be *predicative relative to a variable x* if and only if it maps all free variables to values $\leq \sigma(\text{'}x\text{'}) + 1$ and all bound variables to values $\leq \sigma(\text{'}x\text{'})$. A formula φ is said to be *predicatively stratified relative to x* if and only if φ has a predicative stratification relative to x.

The instances of comprehension provided in TTP are the assertions "$\{x \mid \varphi\}$ exists" such that φ is predicatively stratified relative to x.

There is a less restrictive type theory TTI which we have elsewhere described as "mildly impredicative". In this theory, we are permitted to "define" type n objects using quantification over all of type n, but we are still not permitted to mention objects of type higher than n in the definition of a type n object. The definition of the set of natural numbers succeeds in TTI but fails in TTP (that is, the comprehension axiom for $\mathcal{N}$ (in appropriate types) is acceptable in TTI but not in TTP). There is a fragment of NF obtained by dropping types from axioms of TTI: this theory NFI has the extensionality axiom of NF and all comprehension axioms "$\{x \mid \varphi\}$ exists" where there is a stratification σ of φ such that the range of σ includes no value greater than $\sigma(\text{`}x\text{'}) + 1$. We provide a suitable definition.

Definition 3.2 A stratification σ of a formula φ is said to be *mildly impredicative relative to a variable* x if and only if it maps all variables to values $\leq \sigma(\text{`}x\text{'}) + 1$. A formula φ is said to be *mildly impredicatively stratified relative to* x if and only if there is a mildly impredicative stratification of φ relative to x.

Then the comprehension scheme of NFI provides for the existence of $\{x \mid \varphi\}$ just in case φ is mildly impredicatively stratified relative to x.

Denote the singleton operation by ι: $\iota x = \{x\}$. Let $\iota^k x$ represent the k-fold iterated singleton of x: for example, $\iota^2 x = \{\{x\}\}$. (It is useful to note for the sequel that we use the notation ι^k"x for the set of k-fold singletons of elements of x.) For any stratified formula φ, the set which we can informally describe as $\{\iota^k x \mid \varphi\}$ is provided by an instance of stratified comprehension of the form $\{y \mid (\exists x.\text{“}y = \iota^k x\text{”} \wedge \varphi)\}$; the type of y will be k higher than the type of x, and if k is large enough the formula $(\exists x.\text{“}y = \iota^k x\text{”} \wedge \varphi)$ will be predicatively stratified relative to y, so the existence of the set will be asserted by TTP. (We assume that the reader can provide the formal definition of the sequence of formulas abbreviated "$y = \iota^k x$" if he or she really wants it). This is sufficient to show that the axiom of set union adjoined to TTP (or to TTI) gives a theory equivalent to TT; application of set union k times to $\{\iota^k x \mid \varphi\}$ yields $\{x \mid \varphi\}$. For the same reason, adjoining the axiom of set union to NFP or NFI gives a theory equivalent to NF.

Theorem 3.3 (Crabbé) *NFP (and so NFI) proves the axiom of infinity.*

Proof: One may define each concrete Frege natural number in NFP just as in NF; moreover, one may define the successor operation on Frege natural numbers just as in NF. We give the explicit definitions.

Definition 3.4 0 is defined as $\{\varnothing\}$, the set of all sets with no elements.

Definition 3.5 For any set A, $A + 1$ is defined as $\{a \cup \{x\} \mid a \in A \wedge x \notin a\}$, the set of all disjoint unions of elements of A with singletons.

Observation 3.6 The sets 0 and $A + 1$ are provided by instances of predicative stratified comprehension.

Definition 3.7 We call a set I *inductive* if and only if $0 \in I$ and $(\forall a.a \in I \rightarrow a + 1 \in I)$.

Definition 3.8 We say that an object n is a *natural number* if and only if it belongs to all inductive sets. We may refer to the class of natural numbers as $\mathcal{N}$, but we do not assume that $\mathcal{N}$ is a set.

Observation 3.9 The set of all inductive sets is predicatively defined (though this is not important to us here); the set of all natural numbers is *not* provided by the scheme of predicative stratified comprehension (though it is provided by the scheme of "mildly impredicative" stratified comprehension of NFI) and it turns out that there are models of NFP in which $\mathcal{N}$ is not a set.

Definition 3.10 A set is said to be *finite* if and only if it belongs to some natural number. There is no presumption that the finite sets make up a set. A set is said to be *infinite* if it is not finite. We refer to the assertion that the universal set V is infinite as the "axiom of infinity".

We define a set $U = \{A \mid (\forall a \in A.(\forall b \subseteq a.(\exists c.(\forall d.d \in c \equiv (\exists e.d \in e \wedge e \in b)))))\}$; informally, this is the set of all sets A such that all subsets of elements a of A have unions. We provide it in fully expanded form to allow the reader to check that predicative stratified comprehension does provide this set in spite of the fact that the impredicative notion of set union is involved; the relative type of each variable in the set definition is less than or equal to that of A.

We show that U is inductive. It is necessary to check that if $A \in U$, $a \in A$, and $x \notin a$, then all subsets of $a \cup \{x\}$ have unions. If such a subset b does not contain x, it has a union because it is a subset of a. If it contains x, then $b - \{x\}$ has a union because it is a subset of a, and $x \cup \bigcup(b - \{x\})$ exists and equals $\bigcup b$. Thus the set U is inductive and thus contains all natural numbers.

If the universe V were finite, it would belong to some natural number, which would in turn belong to U. It would follow from this that every subset of V, and so every set, would have a union, which implies that all consequences of full NF, including the assertion "V is infinite" would hold. This is a contradiction, so V must be infinite.

Other forms of the axiom of infinity given in the NF literature, such as the assertion that the empty set is not a natural number, can be proved. It is easy to see that if $\varnothing$ belongs to all sets containing 0 and closed under successor, then so must ιV, the singleton of the universe, because $\varnothing$ itself and ιV are the only sets with $\varnothing$ as their successor. But we have just shown that ιV cannot be a natural number, because no set containing V can be a natural number. Thus the empty set cannot be a natural number either. The proof is complete. □

In [8], Quine presented a definition of the ordered pair suitable for use in TT with Infinity or NF which is "type level": the pair is of the same type as its projections. Note that this is not true of the Kuratowski pair: $\{\{x\}, \{x, y\}\}$ is two types higher than its projections x and y.

We describe Quine's definition of the ordered pair. Define a map σ as follows: $\sigma(n) = n + 1$ for $n \in \mathcal{N}$ and $\sigma(x) = x$ for $x \notin \mathcal{N}$. Define $\sigma_1(A)$ as $\sigma``A$ for any set $A$, and define $\sigma_2(A)$ as $(\sigma``A) \cup \{0\}$. Then the Quine pair $\langle A, B\rangle$ is defined as $(\sigma_1``A) \cup (\sigma_2``B)$. Let $\langle A, B\rangle^-$ be the set of elements of $\langle A, B\rangle$ which do not contain 0 and let $\langle A, B\rangle^+$ be the set of elements of $\langle A, B\rangle$ which do contain 0. Define $\sigma_3(A)$

as σ^{-1}"A for any set A. It is straightforward to establish that σ_3"$(\langle A, B\rangle^-) = A$ and σ_3"$(\langle A, B\rangle^+) = B$, so we can recover the projections from the pair.

Adapting this definition of the pair to TTP or NFP requires care. The difficulty is that the class of natural numbers is not necessarily a set in the predicative theories, so there is no reason to believe that σ_1, σ_2, or σ_3 are operations which take sets to sets. The definition of the pair and its projections is otherwise predicatively acceptable.

A prerequisite for the modified definition is the notion of a cardinal which has to be defined in terms of the usual Kuratowski pair at this point.

Definition 3.11 The *cardinal of a set* A, written $|A|$, is defined as the set $\{\iota^2 B \mid (\exists f.\ f$ is a bijection from A to $B)\}$, where the notion of bijection (and the notion of function on which it depends) are defined in the natural way using the Kuratowski pair. The appearance of the double singleton of B as the generic element is due to the fact that the type of the bijection f witnessing the equinumerousness of B and A is two higher than the type of B, and the type of an element of the set being defined must be at least this high to satisfy predicativity restrictions.

It is demonstrable in TTP that cardinals with the same successor are the same cardinal (over an arbitrary type). We can then define the Quine pair in TTP or NFP just as above, except that the map σ will be the successor map on cardinals instead of the successor map on natural numbers. Just as in TT, Infinity needs to hold for the definition to succeed.

The type differential here is considerable: if type n satisfies Infinity (so the definition works), the type of any set A of type n objects is $n+1$, the type of the cardinality $|A|$ of A is $n+4$, the type of the set of cardinals of sets of type n objects is $n+5$, and the type in which the pair is defined is $n+6$. The pair is also definable on each type above $n+6$ (all types above an infinite type are infinite; types below an infinite type need not be (internally) infinite in predicative type theory).

Since a type-level pair is available in NFP, functions and relations can be defined to be one type higher than the elements of their domains and ranges instead of three types higher. One consequence of this is that the bizarre definition of cardinals as sets of double singletons forced on us by predicativity restrictions can be replaced with the usual one ($|A| = \{B \mid$ there is a bijection from A onto $B\}$) once the type-level pair is available.

It is interesting to observe that the definitions of equinumerousness using the Kuratowski pair and the modified Quine pair do not coincide. The precise relationship is that sets A and B have a bijection between them represented as a set of Kuratowski pairs precisely if ι^2"A and ι^2"B have a bijection between them represented as a set of modified Quine pairs. In the predicative context, it is possible for the singleton images of two sets to have the same cardinality when the sets themselves have different cardinalities.

We present a proof of the consistency of NFI (and thus of NFP). This proof is not the same as the one in [1]; it is more closely related to the consistency proof for NFU given by Jensen.

In preparation for this proof, we cite basic results of Specker (from [11]) which we will not prove here.

Theorem 3.12 (Specker) *NF (respectively, NFU, NFP) is consistent if and only* *if TT (respectively, TTU, TTP) with the ambiguity scheme is consistent.*

Theorem 3.13 (Specker) *NF (respectively, NFU, NFP) is consistent if and only if there is a model of TT (respectively, TTU, TTP) which is isomorphic to its submodel obtained by relabeling each positive type $n+1$ as type n.*

Specker's original results were for NF alone (the other theories had not yet been proposed) but they adapt easily to the subtheories.

It is straightforward to construct a model of TTI in which all infinite sets in each type are the same size. We describe an inductive construction of such a model. Let type 0 be any countably infinite set on which a notion of ordered pair is defined. Suppose that types 0-n have been defined with the following properties: each type is countably infinite and supports a type level ordered pair, and each type $i+1$ contains a bijection between type i and the set of i-fold singletons of type 0 objects. We provisionally let type $n+1$ be the true power set of type n. We define the pair in type $n+1$ as follows: let x and y be two distinct elements of type n and let (a, b) be $\{(a', x) \mid a' \in a\} \cup \{(b', y) \mid b' \in b\}$ for any type $n+1$ objects a and b. The problem we have is that the provisional type $n+1$ is too large (it is uncountably infinite). The solution is to replace the provisional model of types $0 - n+1$ by a Skolem hull of the theory of the provisional model with equality, membership, and pair projections in each type as predicates, and with all the elements of types $0 - n$ as constants. This will be a countably infinite structure in which types $0 - n$ are unchanged and the new type $n+1$ satisfies the desired conditions.

In such a model, any set of a given type which is externally seen to be infinite has the same cardinality in the internal sense of the model as any other such set; all externally infinite sets of a given type are externally countably infinite, so the provisional version of the next higher type contains maps witnessing their equinumerousness, and the formation of the Skolem hull preserves the existence of such maps, witnessing the fact that the sets have the same cardinality internally.

It is thus possible to define a "membership" relation of objects of type i in objects of any type $j > i$ (not just $j = i+1$). This relation $\in_{ij}$ is defined as follows: for each $j > i$ choose a bijection f_{ij} from ι^{j-i-1}"V^i (where V^i is notation for type i as a set) to V^{j-1}, and define $x \in_{ij} y$ as $f_{ij}(\iota^{j-i-1}x) \in y$. It is straightforward to establish that any increasing sequence of types in the given model is a model of TTI when appropriate $\in_{ij}$'s are used as membership relations between successive types in the sequence.

With each finite set of sentences Σ of the language of TT involving types 0 to $n-1$, we associate a partition of the n-element sets of types of our model. The partition will have no more than $2^{|\Sigma|}$ compartments: it will be determined by the truth values of the sentences of Σ when the types 0 to $n-1$ are replaced by the elements of the n-element set in increasing order, with appropriate replacements of occurrences of $\in$ with $\in_{ij}$'s.

By Ramsey's theorem, there is an infinite sequence of types which is typically ambiguous for the partition associated with Σ and so satisfies typical ambiguity for the formulas in Σ. By compactness, the scheme of typical ambiguity can consistently be adjoined to TTI. By Specker's theorems cited above, NFI is equiconsistent with TTI with the ambiguity scheme. Since NFP is a fragment of NFI, NFP is also consistent.

If we work in a nonstandard model of set theory with an external automorphism j such that $j(N) > N$ for a (nonstandard) natural number N, we can build an explicit model of NFI. The infinite sequence of types $\{j^i(N)\}_{i\in\mathcal{N}}$ (a proper class in our nonstandard model of set theory) in the model of TTI constructed above, with the membership relations $\in_{j^i(N)j^{i+1}(N)}$ between successive types, gives a model of TTI with an isomorphism between itself and its submodel obtained by relabeling each positive type $n+1$ as type n. A model of NFI is obtained as follows: use type N as the domain of the model and define $x \in_{NFI} y$ as $x \in_{Nj(N)} j(y)$.

***4 The ramified theory of types,* TTP *and* NFP** We introduce a formalization (with variants) of the ramified theory of types of the second edition of Russell and Whitehead's *Principia Mathematica*.

The guiding idea of the ramified theory of types is that no object may be defined using quantification over a totality to which the object itself belongs. The variations in our treatment stem from whether one wishes to include types of relations or just types of sets. We will start with a formalization which admits just types for sets (following Crabbé in [2]) and indicate how it could be adapted to handle types of relations (or even functions).

A type in the scheme we adopt following Crabbé is a finite strictly increasing sequence of natural numbers whose first term is 0. The sequence whose first and only term is 0, which we will denote by (0), will be the base type, corresponding to type 0 in TT. We use the notation s^- for the sequence obtained by dropping the last term of the sequence s (when $s \neq (0)$). We use the notation $\max(s)$ for the last term in the (increasing) sequence s.

The elements of an object of type s will be objects of type s^-. The extensionality axiom of the ramified theory of types asserts as usual that objects of any type other than the base type are equal if and only if they have the same elements. The comprehension scheme provides for the existence of $\{x^{s^-} \mid \varphi\}^s$, the set of type s of all x of type s^- such that φ, just in case φ has no parameter of a type t such that $\max(t) > \max(s)$ and contains no bound variable of a type t such that $\max(t) \geq \max(s)$. This criterion should remind one of the criteria for comprehension in TTP.

This scheme has the effect that any set can be defined by quantification only over types represented by sequences with a smaller maximum element and using parameters of types represented by sequences with no larger maximum element, which has the effect of enforcing predicativity. It appears to be more expressive than TTP because it allows one to define subsets of any given type using quantifiers over any other type, by providing each type with multiple power set types. Metaphorically, one can think of the base type (0) as being "created" first, then the sole type t with $\max(t) = 1$, then the two types with $\max(t) = 2$, and at the nth stage all types t with $\max(t) = n$ would be created: (0) is the type of individuals created first; $(0\,1)$ is the type of sets of individuals created at the second stage; $(0\,1\,2)$ is the type of sets of type $(0\,1)$ objects (sets of sets of individuals) created at the third stage, while $(0\,2)$ is a new type of sets of individuals, also created at the third stage; at the next stage, $(0\,3)$ will be yet another type inhabited by sets of individuals, while $(0\,1\,3)$ and $(0\,2\,3)$ will be two subtly different new types of sets of sets of individuals, inhabitants of which will have elements of types $(0\,1)$ and $(0\,2)$, respectively. The details of the restrictions on comprehension are motivated in the same way as in TTP: in the definition of a set of a certain type, quantification over types not yet created is obviously not possible; parameters of types created at the same stage are permitted because we do not imagine that all objects in any given type are created simultaneously. Each type has a new "power set" created at each subsequent stage.

This already "ramified" picture can be further ramified by allowing relations of arbitrary arity or even functions. The types could then be coded as nonempty trees labeled with natural numbers and with 0 at all leaves, rather than sequences of positive natural numbers: the code for an n-ary relation type would have n ordered immediate subtrees, the ith ordered subtree being the type of the ith argument to be supplied to the relation. The code for a function type would have an additional immediate subtree for the type of its output. The trees would be "increasing" in the sense that the natural number labeling the root of a tree would be larger than the natural numbers labeling the roots of each child tree. The extensionality and comprehension schemes would be essentially the same, with the value of the function coding a type at the root of its domain playing the role of $\max(s)$ above.

The ramified type system, even when augmented with relations and functions, is mutually interpretable with the much simpler predicative type theory TTP. An immediate corollary is that the ramified theory of types (with the assumption that the base type is infinite) is interpretable in the very simple set theory NFP. Of course, there are some philosophical questions about the status of NFP, which is not in any obvious sense a predicative theory.

We will now establish the mutual interpretability of TTP and the set version of ramified type theory. It is straightforward to see that ramified type theory interprets TTP. In fact, there are many interpretations: any sequence t_i of types in ramified type theory with the property that $t_i = t_{i+1}^-$ for each index i provides a direct interpretation of TTP, using the type t_i as type i.

The subtler thing is to see that the apparently weaker theory TTP actually has the full expressive power of ramified type theory. We describe the interpretation, then verify that it works. The base type (0) of ramified type theory is coded by type 0. Each type t of ramified type theory is coded by a subset of the type $\max(t)$ of TTP (a set of type $\max(t) + 1$). This might appear to create a problem with membership, as $\max(s^-)$ is not necessarily the predecessor of $\max(s)$; this is handled by the stipulation that membership of elements of type s^- in elements of type s is coded by membership of suitably iterated singletons of the elements of the interpreted type $\max(s^-)$ in objects of the interpreted type $\max(s)$. If type s^- is coded by a subset A of type $\max(s^-)$, type s is coded by $\mathcal{P}\iota^{\max(s)-\max(s^-)-1}$"$A$, the set of all sets of $(\max(s) - \max(s^-) - 1)$-fold iterated singletons of elements of A. This is enough to determine the representation of each type of the ramified theory precisely.

In the case of the more complex theories with function and relation types, the singleton operation will need to be applied a different number of times to each argument of the function or relation when the arguments are of different types: for example, a stage 4 type inhabited by relations between a stage 1 type and a stage 2 type will be represented by a set of type 4 objects whose elements are pairs with first projection a double singleton of a type 1 object and second projection a singleton of a type 2 object. The theories with relation and function types can only be handled in the way we describe if a type-level ordered pair is available; otherwise the type of an *n*-tuple relative to its arguments will depend on *n*, making a uniform representation impossible. It is necessary in this case to add the type-level ordered pair as a primitive notion to TTP. The pair is only definable in general models of TTP in types 6 and above; truncating a model of TTP at type 6 gives a model of TTP in which a type level pair is available at every type, so we can see that we do not essentially strengthen TTP by adding the type-level pair as a primitive.

The verification of extensionality in this interpretation is trivial. The verification of comprehension is hardly less trivial when it is recognized that the characteristic $\max(s)$ of types *s* which controls the predicativity restriction in the ramified theory is mapped precisely to the type in TTP, and the restrictions on these parameters in comprehension in the two theories are the same: a permitted comprehension axiom in the ramified theory will map to a permitted comprehension axiom in TTP. A quantifier over a type *s* of ramified type theory will be interpreted by a quantifier over the type $\max(s)$ of TTP bounded by a set of type $\max(s)+1$, which will become an additional parameter of type $\max(s)+1$ in the interpretation. A comprehension axiom $\{x \mid \varphi\}$ of the ramified type theory introducing a set of type *s* satisfies the condition that each bound variable will be of a type *t* with $\max(t) < \max(s)$; each bound variable in the interpreted version is thus of type $\max(t) < \max(s)$, bounded by a parameter of type $\max(t)+1 \leq \max(s)$. The interpreted version of the comprehension axiom introduces the set of iterated singletons in type $\max(s)-1$ of objects of type $\max(s^-)$ which satisfy a condition involving bound variables of types $< \max(s)$ and parameters of types $\leq \max(s)$ (the interpreted versions of the original parameters plus parameters introduced to bound interpreted quantifiers); this will be a comprehension axiom of TTP.

So we have completed the demonstration of the following.

Theorem 4.1 *TTP and ramified type theory are mutually interpretable; ramified type theory is interpretable in NFP.*

Since TTP interprets ramified type theory, it follows immediately that NFP interprets ramified type theory. NFP is a very appealing theory; it is not very strong (Crabbé showed that it is actually weaker than Peano Arithmetic) but it is very expressive, and it is strong enough to do a good deal of elementary mathematics. It combines Russell's concerns about predicativity with Quine's solution to the inelegance of types and the "hall of mirrors" effect of typical ambiguity. In addition, there is the surprise that NFP actually proves infinity. But there is the potential philosophical difficulty that NFP (in spite of its etymology) may not really be a predicative theory; sets can belong to themselves in this theory after all. We do not rule out the possibility that there is a justification for NFP on predicative grounds, but we haven't yet produced one ourselves.

***5* NFI *has the same strength as* PA$_2$** NFI is evidently stronger than NFP. In NFI, the set of natural numbers exists, and in fact NFI interprets PA_2 (second-order arithmetic) in the natural way; the theory of natural numbers and sets of natural numbers defined in the natural way in NFI provides a direct interpretation of PA_2. Comprehension axioms defining sets of natural numbers in NFI may freely employ quantifiers over natural numbers (one type lower than the type of the set being defined) and over sets of natural numbers (at the same type as the set being defined); this is all that is needed to interpret comprehension axioms of PA_2.

The remainder of this section is devoted to the proof of the much less obvious fact that the consistency of PA_2 implies the consistency of NFI: the two systems have exactly the same consistency strength.

We first define quantifier classes of formulas in a standard way. In the following discussion, all formulas are assumed to be presented in prenex form with all quantifiers over sets of natural numbers preceding all quantifiers over natural numbers. (A quantifier over natural numbers appearing before a quantifier over classes could be converted to a quantifier over one-element sets of natural numbers; it can also be absorbed into a following quantifier over sets of natural numbers by considering the fact that a function from a finite list of natural numbers to a set of natural numbers can be coded as a set of natural numbers).

A Σ_0 or Π_0 formula is a formula with no quantifiers over sets of natural numbers at all (the two terms are synonymous). A Σ_{n+1} formula results when one or more existential quantifiers are prefixed to a Π_n formula; a Π_{n+1} formula results when one or more universal quantifiers are prefixed to a Σ_n formula. It is useful to note that any finite tuple of sets of natural numbers can be coded by a single set of natural numbers, so one can think of a Σ_n formula as having a single existential quantifier leading a string of n alternating quantifiers over sets of natural numbers, and a Π_n formula as having a single universal quantifier leading a string of n alternating quantifiers over sets of natural numbers.

Satisfaction of formulas and truth of sentences can be defined in the usual way in PA_2 (for restricted classes of sentences, of course). It is not hard to show that satisfaction for formulas without quantifiers over sets (Σ_0 sentences) is Σ_1. The reader may consult Rogers [9] for a proof of the fact that satisfaction of Σ_{n+1} formulas can be represented by a Σ_{n+1} formula, which is what we require for this development. It is, of course, impossible (by Tarski's well-known theorem on the definability of truth) for PA_2 to represent satisfaction of formulas of PA_2 in general. For one who is familiar with the details of the proof of Tarski's theorem, it might be useful to point out that the fact that the class of Σ_{n+1} sentences can contain a representation of truth for sentences of that same class is not problematic because this class is not closed under negation.

We represent by TTI_k the fragment of TTI axiomatized by those axioms of TTI which are Σ_i sentences (or which have prenex forms which are Σ_i sentences) for some $i \leq k$. We show that TTI_k is interpretable in PA_2. We do this by showing that TTI_k^n, the n-type fragment of TTI_k, is interpretable for each natural number n in a uniform manner. Clearly TTI_k^1 is interpretable in PA_2; all that is required is a domain of elements with the equality relation on them; we stipulate further that we will use a countably infinite domain as our type 0.

Suppose that we have modeled TTI_k^{n+1} (whose highest type is type n); each of the types $0 - n$ will be a countably infinite structure whose elements are natural numbers. We show how to model TTI_k^{n+2}. We provisionally take our type $n+1$ to consist of the subsets of the type n of our model; this must be provisional because this makes type $n+1$ too large with elements of the wrong sort. We can define the Σ_k

theory of this structure using the definition of truth for Σ_k sentences. We then enrich the language of TTI_k^{n+2} with Hilbert symbols (so that we can represent witnesses to existential statements) and with all elements of types $0-n$ as constants. We define a maximal consistent extension of the Σ_k theory of our provisional model of types $0-n+1$ with the enriched language including Hilbert symbols; since our language is countable, we can do this without an appeal to the axiom of choice. From this maximal consistent extension, we can extract a countably infinite term model of types $0-n+1$, satisfying TTI_k^{n+2}. This construction can be described in a uniform manner (using a higher quantifier class than Σ_k!); so we can construct a model of TTI_k in PA_2. We can then follow the proof of consistency of NFI above; the infinite Ramsey theorem is provable in PA_2, so we find that PA_2 proves the consistency of NFI_k for each concrete k. This means that we can actually build models of NFI_k in PA_2 (as, for instance, by the method of taking maximal consistent extensions of theories to build a term model).

Since PA_2 proves the consistency of each NFI_k, it must be the case that the consistency of PA_2 implies the consistency of NFI (which is the union of all the NFI_ks). This also demonstrates that NFI cannot be finitely axiomatized, as otherwise NFI would be equivalent to some NFI_K and PA_2 would prove the consistency of this NFI_K, so of NFI, so of PA_2 itself, which is impossible by Gödel's second incompleteness theorem. It is known that NF and NFP can be finitely axiomatized. The standard reference for a finite axiomatization of NF is Hailperin [3] but the axiomatization given there cannot easily be adapted to NFP. The finite axiomatization for NFU given in Holmes [5] can be adapted to NFP, if one adds strong extensionality and drops the axiom of set union.

So we have completed the proof of the following.

Theorem 5.1 *PA_2 is equiconsistent with NFI.*

6 Conclusions and questions We believe that these results basically settle the status of the theory NFI. Crabbé showed that its consistency strength was less than that of third-order arithmetic; here we show that it is exactly that of second-order arithmetic. The apparent ability to apply a "power set" construction repeatedly to an infinite set in NFI is in some sense illusory: NFI can be encoded in a system in which the power set operation is applied to an infinite set just once. But the apparent ability to iterate the power set construction gives NFI expressive power greater than that of second-order arithmetic in practical terms: in NFI we can define the rational numbers just as we would in NF or ZFC, then define the reals as Dedekind cuts in the rationals, and go on to represent such things as functions from the reals to the reals as well. The analysis one is then able to do is predicative; the least upper bound property does not hold in general for these "reals", but the expressive power of a full set theory is convenient nonetheless.

We have shown elsewhere (e.g., in Holmes [4]) that NFI admits consistent extensions of arbitrarily high consistency strength, including ones in which arithmetic and analysis are standard. Some questions remain about NFP. We are interested in finding out its exact consistency strength and exploring the extent to which elementary mathematical constructions can be carried out in NFP. Even more than NFI,

NFP appeals to us as combining extreme weakness in terms of consistency strength with the full expressive power of set theory. We are also interested in the (philosophical rather than mathematical) question of whether NFP can really be motivated on a purely predicative basis: the fact that it seems to be no stronger than TTP + Infinity suggests that it may be possible to do this.

REFERENCES

[1] Crabbé, M., "On the consistency of an impredicative subsystem of Quine's NF," *The Journal of Symbolic Logic*, vol. 47 (1982), pp. 131–36. Zbl 0487.03024 MR 83d:03061 1, 1, 3

[2] Crabbé, M., "Ramification et prédicativité," *Logique et Analyse*, vol. 84 (1978), pp. 399–419. Zbl 0449.03009 MR 81e:03006 1, 4

[3] Hailperin, T., "A set of axioms for logic," *The Journal of Symbolic Logic*, vol. 9 (1944), pp. 1–19. Zbl 0060.02201 MR 5:197a 5

[4] Holmes, M. R., "The equivalence of NF-style set theories with 'tangled' type theories; the construction of omega-models of predicative NF (and more)," *The Journal of Symbolic Logic*, vol. 60 (1995), pp. 178–89. Zbl 0819.03044 6

[5] Holmes, M. R., *Elementary Set Theory with a Universal Set*, Academia, Louvain-la-Neuve, 1998. Zbl 0959.03001 MR 2001c:03089 5

[6] Jensen, R. B., "On the consistency of a slight(?) modification of Quine's NF," *Synthese*, vol. 19 (1969), pp. 250–63. Zbl 0202.01001 2

[7] Quine, W. V., "New foundations for mathematical logic," *American Mathematical Monthly*, vol. 44 (1937), pp. 70–80. Zbl 0016.19301 2

[8] Quine, W. V., "On ordered pairs," *The Journal of Symbolic Logic*, vol. 10 (1945), pp. 95–96. Zbl 0060.12414 MR 7:185c 3

[9] Rogers, H., *The Theory of Recursive Functions and Effective Computability*, The MIT Press, Cambridge, 1988. See Theorem XIII, Corollary XIIIa, Corollary XIV, in Section 16.2, pp. 386–90. 5

[10] Specker, E. P., "The axiom of choice in Quine's 'New foundations for mathematical logic'," *Proceedings of the National Academy of Sciences of the USA*, vol. 39 (1953), pp. 972–75. Zbl 0051.03705 MR 15:493b 2

[11] Specker, E. P., "Typical ambiguity," pp. 116–124 in *Logic, Methodology and Philosophy of Science*, edited by E. Nagel, Stanford University Press, Palo Alto, 1962. Zbl 0156.02101 MR 28:1129 3

Department of Mathematics
Boise State University
1910 University Drive
Boise ID 83725
email: holmes@math.idbsu.edu