# The usual model construction for *NFU* preserves information

M. Randall Holmes

January 16, 2012

**Abstract**

The usual construction of models of *NFU* (New Foundations with urelements, introduced by Jensen) is due to Maurice Boffa. A Boffa model is obtained from a model of (a fragment of) ZFC with an automorphism which moves a rank: the domain of the Boffa model is a rank that is moved. "Most" elements of the domain of the Boffa model are urelements in terms of the interpreted *NFU*. The main result of this paper is that the restriction of the membership relation of the original model of set theory with automorphism to the domain of the Boffa model is first-order definable in the language of *NFU*. In particular, all information about the extensions in the original model of the urelements of the model of *NFU* is definable in terms of *NFU*. A corollary (answering a question of Thomas Forster) is that the urelements in a Boffa model are not homogeneous.

## 1 Introduction

This paper was originally motivated by a question of Thomas Forster about the urelements in a model of Jensen's modification *NFU* of Quine's New Foundations: is it possible for them to be homogeneous? We answer this question in the negative for a familiar class of models (those obtained from a nonstandard model of a fragment of set theory with an automorphism in a way originally proposed by Maurice Boffa) by showing that the membership relation of the underlying model is first-order definable in terms of the interpreted *NFU*, which is not at all obvious and perhaps surprising.

Forster asked whether there are models of *NFU* in which the urelements are homogeneous. By this we mean that for any formula $\phi(a_1 \ldots a_n)$, in which the $a_i$'s are the only parameters, $(\forall a_1 \ldots a_n.(\forall b_1 \ldots b_n.U(a_1) \wedge \ldots \wedge U(a_n) \wedge U(b_1) \wedge \ldots \wedge U(b_n) \rightarrow \phi(a_1 \ldots a_n) \leftrightarrow \phi(b_1 \ldots b_n)))$ holds, where $U(x)$ means "$x$ is an urelement"). Of course a model of *NF* would be such a model of *NFU*. (It is relatively easy to show that the urelements of any model of *NFU* *are* homogeneous with respect to stratified formulas, though we do not give details here: one uses the permutation techniques adapted to *NF* by Dana Scott in [16] and discussed in the context of *NFU* in [2].)

The author's (incorrect!) response was that of course there are such models, because all the urelements in the "usual" models of *NFU* (described in detail below) are indistinguishable in the suitable sense, because all information about the extensions of the urelements of the model is discarded in the construction.

But it turns out that this is not the case, and for a reason which perhaps has more interest than the answer to the original question (which remains unanswered, and may have some of the flavor of the unsolved problem of the consistency of *NF*). The usual construction of a model of *NFU* starts with a model of (some fragment of) ordinary set theory with an automorphism which moves a rank of the cumulative hierarchy. The domain of the model of *NFU* is a rank in this model moved by the automorphism. Most elements of this rank are treated as urelements in the model of *NFU* and it *appears* that information is being discarded in this process. But it is not: the restriction to the domain of the model of *NFU* of the membership relation of the original model of set theory with automorphism turns out to be first-order definable in the model of *NFU*, which quite incidentally gives a negative answer to the question as to whether the urelements in the model of *NFU* are homogeneous.

## 2 *NFU*

We briefly describe the theory *NFU*, along with the related theories *TST* (simple type theory) and *NF* (New Foundations).

Simple type theory *TST* is a strongly streamlined version of the type theory of Russell and Whitehead ([20]). One follows Ramsey in eliminating the orders in [20] ([14]). One follows Wiener ([18]) in noting that since there is a definable ordered pair in typed set theory, one does not need relation types. Theories of this kind were apparently first proposed about 1930 (see

[19] for historical remarks).

*TST* is a first-order theory with sorts (called types) indexed by the natural numbers. The typing conditions for atomic formulas are briefly given by the templates $x^n = y^n$, $x^n \in y^{n+1}$. The axioms of *TST* are

**Extensionality:**

$$(\forall A^{n+1} B^{n+1}.A^{n+1} = B^{n+1} \leftrightarrow (\forall x^n.x^n \in A^{n+1} \leftrightarrow x^n \in B^{n+1}))$$

is an axiom for each $n$.

**Comprehension:** For each formula $\phi$ in which $A^{n+1}$ is not free, and each variable $x^n$,

$$(\exists A^{n+1}.(\forall x^n.x^n \in A^{n+1} \leftrightarrow \phi))$$

is an axiom.

Axioms of Infinity and Choice are usually adjoined, whose exact form need not be investigated at this point.

In [13], W. v. O. Quine proposed that the sorts of *TST* could be collapsed so as to obtain an unsorted first-order theory with equality and membership whose axioms are exactly the axioms of *TST* with distinctions of type between variables ignored (in a way which does not introduce any identification between variables of different types). This theory is called *NF* (New Foundations, after the name of the paper [13]). It is traditional to present it in a way which does not depend on the definition of another theory:

**Extensionality:**

$$(\forall AB.A = B \leftrightarrow (\forall x.x \in A \leftrightarrow x \in B))$$

is an axiom.

**Definition:** Let $\phi$ be a formula in the language of first-order logic with equality and membership. We say that a function $\sigma$ from variables to natural numbers is a *stratification* of $\phi$ iff for each atomic subformula '$x = y$' of $\phi$ we have $\sigma('x') = \sigma('y')$ and for each atomic subformula '$x \in y$' we have $\sigma('x') + 1 = \sigma('y')$. We say that $\phi$ is *stratified* iff there is a stratification of $\phi$.

**Stratified Comprehension:** For each stratified formula $\phi$ in which $A$ is not free, and each variable $x$,

$$(\exists A.(\forall x.x \in A \leftrightarrow \phi))$$

is an axiom.

It should be clear that this is an equivalent description of the theory, as the stratified formulas are exactly those which can be obtained from well-formed formulas of *TST* by dropping type distinctions without introducing identifications between variables.

In fact, the notion of stratification can be eliminated from the definition of the theory and with it the last reference to even relative notions of type: the axiom scheme of stratified comprehension is equivalent to the conjunction of finitely many of its instances. The usually referenced though far from the nicest presentation of this is in [5]. A nicer presentation is found in the author's [8], though this needs to be modified to eliminate references to a primitive ordered pair (the way to do this is indicated there).

The consistency of *NF* remains an open question. In 1969 ([9]) R. B. Jensen proved the consistency of *NFU* (New Foundations with urelements), which differs from *NF* only in its formulation of Extensionality:

**Weak Extensionality:**

$$(\forall AB.(\exists y.y \in A) \rightarrow A = B \leftrightarrow (\forall x.x \in A \leftrightarrow x \in B))$$

is an axiom.

Notice that *NFU* has the same comprehension scheme as *NF*. The idea is that *NFU* allows the existence of many urelements without elements in addition to the empty set. Strictly speaking, the formulation above does not allow one to specify an empty set. An inessential and very convenient modification of the theory is to add the empty set as a primitive (which allows one further to define a sethood predicate). We present a full axiomatization of *NFU* with empty set as a primitive notion:

**Empty Set:** $(\forall x.x \notin \emptyset)$

**Definition:**

$$\texttt{set}(x) \equiv_{\texttt{def}} x = \emptyset \vee (\exists y.y \in x)$$

**Weak Extensionality:**

$$(\forall AB.\mathtt{set}(A) \wedge \mathtt{set}(B) \rightarrow A = B \leftrightarrow (\forall x.x \in A \leftrightarrow x \in B))$$

is an axiom for each $n$.

**Stratified Comprehension:** For each stratified formula $\phi$ in which $A$ is not free, and each variable $x$,

$$(\exists A.\mathtt{set}(A) \wedge (\forall x.x \in A \leftrightarrow \phi))$$

is an axiom.

Specker showed in [17], 1953 that *NF* disproves Choice (and so proves Infinity). Jensen showed that *NFU* is consistent with Infinity and Choice, and further that *NFU* is consistent with the negation of Infinity.

# 3 The Boffa model construction

We do not describe the original consistency proof of *NFU* due to Jensen. Instead, we give a related model construction due to Maurice Boffa (in [1]). This is appropriate as the aim of this paper is to prove a result about models of the sort described by Boffa (the class of models we consider is actually slightly less general than the class of models discussed by Boffa).

We use the term "Mac Lane set theory" for Zermelo set theory with separation restricted to bounded formulas. This system was proposed as a foundation for mathematics by Saunders Mac Lane in [10]. Let $M$ be a model of Mac Lane set theory with the additional axiom that every set belongs to a rank $V_\alpha$. The additional axiom provides no essential additional strength (see [11]). We can further suppose, using the Ehrenfeucht-Mostowski theorem of model theory ([3]), that $M$ has an external automorphism $j$ that moves a "rank" $V_\alpha$ (and of course its "ordinal index" $\alpha$, which is not a standard ordinal). We can further suppose without loss of generality that $M \models \alpha > j(\alpha)$ (as we could replace $j$ with $j^{-1}$ to make this true in the worst case). We stipulate that the structure we are talking about is of the form $\langle M, \in_M, j, \alpha \rangle$, so that we can talk about the automorphism. The structure $\langle M, \in_M, j, \alpha \rangle$ is a model of Mac Lane set theory + "every set belongs to a rank" with the following modifications: $j$ and $\alpha$ are added to the language, along with axioms asserting that $\alpha$ is an infinite ordinal, $V_\alpha$ exists, and $j(\alpha) < \alpha$, and with Separation restricted to formulas in which $j$ does not appear (which can be extended to formulas in which $j$ appears in parameters).

We construct a Boffa model $\langle B, \in_B, j \rangle$ of *NFU* (with an additional function symbol $j$ which cannot appear in instances of stratified comprehension, except in parameters; we include it merely so that $j$ makes sense in $B$-formulas) as follows. $B = \{x \mid M \models x \in V_\alpha\}$: $B$ is the collection of $M$-elements of $V_\alpha$. $x \in_B y$ is defined as $j(x) \in_M y \wedge y \in_M V_{j(\alpha)+1}$. Of course $B \models x = j(y)$ iff $M \models x = j(y) \wedge y \in V_\alpha$. Notice that if $M \models u \in V_\alpha - V_{\alpha+1}$ then $B \models (\forall x.x \notin u)$ (the empty set of $M$ is also elementless in the model of *NFU*, and naturally taken to be the empty set of $B$). The $M$-elements of $V_\alpha - V_{\alpha+1}$ are the urelements of the model $B$. It should be evident that weak extensionality holds in $B$.

We outline the proof that Stratified Comprehension holds in $B$ (following [4]). Let $\phi$ be a stratified formula in the language of *NFU*. It translates directly into a formula $\phi_1$ in the language of $M$ (by replacing each $u = v$ with $u = v$, each $\emptyset$ with $\emptyset^M$, and each $u \in v$ with $j(u) \in v \wedge v \in V_{j(\alpha)+1}$. If $M \models$"$\{x \in V_\alpha \mid \phi_1\}$ exists", then we give the name $A$ to the object such that $M \models A = \{x \in V_\alpha \mid \phi_1\}$. We claim that if $A$ exists, then we have $B \models x \in j(A)$ iff $\phi$: $B \models x \in j(A)$ is equivalent to $M \models j(x) \in j(A) \wedge j(A) \in V_{j(\alpha)+1}$, which is equivalent to $M \models j(x) \in j(A)$ because $M \models j(A) \in V_{j(\alpha)+1}$ is true (as obviously $M \models A \in V_{\alpha+1}$), and of course this is equivalent to $M \models x \in A$.

It remains to show that $A$ exists. This does not follow from the separation axiom of $M$ immediately, because the formula $\phi_1$ will usually contain what seem to be essential occurrences of $j$. We show how to convert $\phi_1$ to an equivalent formula in which occurrences of $j$ are confined to parameters, from which it follows that $A$ exists in $M$.

Let $\sigma$ be a stratification of $\phi$. Let $N$ be a constant larger than any element of the range of $\sigma$ (restricted to variables which occur in $\phi$). Replace each atomic formula $u = v$ in $\phi_1$ with the equivalent $j^{N-\sigma(u)}(u) = j^{N-\sigma(u)}(v)$, noting that this is the same as $j^{N-\sigma(u)}(u) = j^{N-\sigma(v)}(v)$ because $\phi$ is stratified. Replace each atomic formula $j(u) \in v$ in $\phi_1$ with the equivalent $j^{N-\sigma(u)}(u) \in j^{N-\sigma(u)-1}(v)$, noting that this is the same as $j^{N-\sigma(u)}(u) \in j^{N-\sigma(v)}(v)$, because $\phi$ is stratified. The remaining atomic formulas $v \in V_{j(\alpha)+1}$ are replaced with $j^{N-\sigma(v)}(v) \in j^{N-\sigma(v)}(V_{j(\alpha)+1})$. The resulting formula we call $\phi_2$. In $\phi_2$, a variable $u$ always occurs in the context $j^{N-\sigma(u)}(u)$, without any further applications of $j$. Now each bound variable $u$ is restricted to $V_\alpha$, so we can replace all occurrences of $j^{N-\sigma(u)}(u)$ with $u$ while replacing the bound on the quantifier with $j^{N-\sigma(u)}(V_\alpha)$. We do this for each bound variable. We also replace $j^{N-\sigma(x)}(x)$ with $x$. We call the resulting formula $\phi_3$. Since $j$ now appears only in parameters in $A_0 = \{x \in j^{N-\sigma(x)}(V_\alpha) \mid \phi_3\}$, this set exists in $M$, and $j^{\sigma(x)-N}(A_0)$ is clearly the desired set $A$. This completes the proof that the model $B$ satisfies Stratified Comprehension.

It is straightforward to establish that $B$ satisfies Infinity and that $B$ satisfies Choice iff $M$ does. If we modified the construction by providing that $M \models$ "$\alpha$ is a finite ordinal", the model $B$ would satisfy *NFU* with the negation of the Axiom of Infinity.

# 4 Recovering information from a Boffa model

If $M \models A \in V_{\alpha+1}$, then for each $y$, $M \models x \in A$ iff $M \models j(x) \in j(A)$ iff $B \models x \in j(A)$: informally, we say that the $M$-set $A$ is implemented by the $B$-set $j(A)$. An unordered pair $\{a, b\}$ is implemented by $\{j(a), j(b)\}$ (the second object being named here in the language of $M$; in the language appropriate to $B$, it would be called $\{a, b\}$: the extremely precise statement of this is that for any $u, a, b$, $B \models u = \{a, b\}$ iff $M \models u = \{j(a), j(b)\}$). The ordered pair $\langle a, b\rangle$ in the sense of $B$ is then $\langle j^2(a), j^2(b)\rangle$ in the sense of $M$. Finally, if $f$ is a function in $M$, we see that $j^3(f)$ is the function with the same extension in $B$: $M \models y = f(x)$ iff $M \models \langle x, y\rangle \in f$ iff $M \models \langle j^2(x), j^2(y)\rangle \in j^2(f)$ iff $B \models \langle x, y\rangle \in j^3(f)$ (clearly $M \models \langle j^2(x), j^2(y)\rangle \in V_{j(\alpha)+1}$). So we see that every $M$-function from $V_\alpha$ to $V_\alpha$ is implemented in $B$ as a function from the universe to the universe (note that $j(V_\alpha)$ is the universal set in the sense of $B$: $M \models x \in V_\alpha$ implies $M \models j(x) \in j(V_\alpha)$ which in turn implies $B \models x \in j(V_\alpha)$. We denote the universal set by $V$ and the collection of all sets by $\mathcal{P}(V)$ in $B$-formulas: the same objects are $j(V_\alpha)$ and $j^2(V_{\alpha+1})$ in $M$.

We introduce a nonstandard piece of notation common in *NF* studies (ultimately derived from [20]), because we will shortly use it. $\iota(x)$ is a notation for $\{x\}$, and so $\iota$"$A$ is notation for the collection of one-element subsets of $A$ (the elementwise image of $A$ under the singleton operation).

We introduce a specific function found in $M$, namely the function $\mathbf{S} = j^3(\{\langle\{x\}, x\rangle \mid x \in j(V_\alpha)\})$. $M \models$ "the domain of $j^{-3}(\mathbf{S})$ is $\iota$"$j(V_\alpha)$ and its range is $j(V_\alpha)$", so of course $M \models j^{-3}(\mathbf{S}) : \iota$"$j(V_\alpha) \to j(V_{\alpha+1})$. An $M$-element of $\iota$"$j(V_\alpha)$ is of the form $\{j(x)\}$, which $B$ sees as $\{x\}$, whereas $M$-elements of $j(V_{\alpha+1})$ are exactly the sets in the sense of $B$: $B \models \mathbf{S} : \iota$"$V \to \mathcal{P}(V)$, i.e., $B$ says that $\mathbf{S}$ is a function from the set of all singletons into the set of all sets. Further, $M \models (\forall x \in V_\alpha . j^{-3}(\mathbf{S})(\{j(x)\}) = j(x))$:

thus $B \models (\forall x.\mathbf{S}(\{x\}) = j(x))$. $\mathbf{S}$, which is a set map, codes the external automorphism $j$ in a certain sense ($j$ is not a set function in either $M$ or $B$).

The function $\mathbf{S}$ has a further property of considerable interest in $B$. Suppose $M \models A \in j(V_{\alpha+1})$ (equivalently, $B \models$ "$A$ is a set"). $B \models x \in A$ iff $M \models j(x) \in A$ iff $M \models j^2(x) \in j(A)$ iff $B \models j(x) \in j(A)$ (certainly $M \models j(A) \in V_{j(\alpha+1)}$) iff $B \models \mathbf{S}(\{x\}) \in \mathbf{S}(\{A\})$. Further, if $B \models y \in j(A)$, then $M \models j(y) \in j(A)$, so $M \models y \in A$, so $M \models y \in V_{j(\alpha)}$, so $M \models y = j(x)$ for some $x$ such that $M \models x \in V_\alpha$, so $B \models y = \mathbf{S}(\{x\})$ for some such $x$. Thus $B \models$ "for any set $A$, $\mathbf{S}(\{A\}) = \{\mathbf{S}(\{x\}) \mid x \in A\}$".

We summarize this statement in the format of an axiom to adjoin to *NFU* (as we did in [7] where we first introduced these ideas with a rather different application in view).

**Axiom of Endomorphism (to adjoin to *NFU*):** There is an injective function $\mathbf{S} : \iota``V \to \mathcal{P}(V)$ such that for any set $A$, $\mathbf{S}(\{A\}) = \{\mathbf{S}(\{x\}) \mid x \in A\}$.

We now observe that in the Boffa model $B$ it is possible to define the restriction of the membership relation of $M$ to $V_\alpha$ in terms of the membership relation of $B$ and the function $\mathbf{S}$: $M \models x \in y$ iff $M \models j(x) \in j(y)$ iff $B \models x \in j(y)$ iff $B \models x \in \mathbf{S}(\{y\})$.

**Definition (*NFU*+ Endomorphism):** $x\,E\,y$ is defined as $x \in \mathbf{S}(\{y\})$. Notice that as the relative types of $x$ and $y$ are the same in the definition of $x\,E\,y$, $E$ is a set relation: $\{\langle x, y\rangle \mid x\,E\,y\}$ exists by stratified comprehension.

$B \models x\,E\,y$ iff $M \models x \in y$.

The final claim which establishes our main result is that if there is a function $\mathbf{S}$ which witnesses the truth of the Axiom of Endomorphism in a Boffa model, then there is exactly one such function. We can then define $x \in^* y$ as "there is a function $\mathbf{S}$ which witnesses the truth of the Axiom of Endomorphism, and $x \in \mathbf{S}(\{y\})$", a statement which can be expanded out into a sentence in the first-order language of *NFU*. Considerations shown above already show that (subject to our final claim) the relation $\in^*$ will coincide with $\in_M$ restricted to the domain of $B$.

Because the domain of $B$ is a rank of the cumulative hierarchy in $M$, there is a notion of ordinal rank of elements of the domain (an element $x$ has rank $\beta$ iff $\beta$ is the least ordinal such that $x \subseteq V_\beta$). It is the case that the ordinal rank of a set $A$ is the successor of the supremum of the set of ranks of elements of $A$. We briefly explain why ranks of the cumulative hierarchy are definable in the weak set theory we are using (recall that this is Mac Lane set theory (Zermelo set theory with bounded separation) with the assumption that every set belongs to a rank, which we will formally state in the course of this development).

**Definition:** A subhierarchy is a well-ordering $\leq_H$ such that the $\leq_H$-least set is $\emptyset$, the $\leq_H$-immediate successor of any set $x$ is $\mathcal{P}(x)$ if it exists, and the $\leq_H$-supremum of a subset $A$ of the domain with no $\leq_H$-maximum is $\bigcup A$. Note that any subhierarchy agrees with the inclusion order on its domain.

**Theorem:** If $\leq_G$ and $\leq_H$ are subhierarchies, either they are equal or one is an initial segment of the other. (proof omitted).

**Definition:** A rank of the cumulative hierarchy is a set which is an element of some subhierarchy.

**Axiom of Rank:** Every set is a subset of some rank (this is a more formal statement of the assumption made about our model $M$ that every set belongs to some rank of the cumulative hierarchy).

**Definition:** The rank of a set $A$ is the smallest rank in the inclusion order containing $A$ as a subset.

**Definition:** The (Scott) order type of a well-ordering $W$ is the set of all well-orderings isomorphic to $W$ and belonging to the smallest rank in the inclusion order containing well-orderings isomorphic to $W$. A (Scott) ordinal is a set which is the Scott order type of some well-ordering.

**Definition:** If $r$ is a rank, we say $r = V_\alpha$ iff $\alpha$ is the Scott order type of the inclusion order on proper subranks of $r$. It is not expected here that $V_\alpha$ exists for every ordinal $\alpha$; it is a consequence of Mac Lane set theory with the Axiom of Rank (and not a consequence of Zermelo or Mac Lane set theory by themselves) that there is an infinite rank, so $V_\omega$ does exist. It can also be noted that if the rank $V_\alpha$ exists, it follows that the usual von Neumann ordinal $\alpha$ exists.

**Definition:** The ordinal rank of a set $x$ is the least $\alpha$ such that $x \subseteq V_\alpha$.

The trick used here to define ordinals in Mac Lane set theory with the Axiom of Rank is due to Dana Scott in [15]. The same trick can be used to represent other sorts of isomorphism class as sets (a common application of Scott's trick is to define cardinals in $ZF$, where the usual von Neumann definition of cardinal number does not work). In Mac Lane or Zermelo set theory without rank there is no reasonable global way to represent cardinal or ordinal number: an extensive development of mathematics in a theory which is essentially Zermelo set theory with the Axiom of Rank, making extensive use of Scott's trick, is found in [12]. The Axiom of Rank does not essentially strengthen Zermelo or Mac Lane set theory; this can be seen in [11].

Suppose that $B$ satisfies the assertion there is another map $\mathbf{S}^*$ such that $\mathbf{S}^* : \iota``V \to \mathcal{P}(V)$ and for any set $A$, $\mathbf{S}^*(\{A\}) = \{\mathbf{S}^*(\{x\}) \mid x \in A\}$. Define $x\, E^*\, y$ as $x \in \mathbf{S}^*(\{y\})$.

Suppose $\mathbf{S} \neq \mathbf{S}^*$. Then there is a minimal ordinal $\beta$ such that there is an element $x$ of $V_\alpha$ of ordinal rank $\beta$ such that $\mathbf{S}(\{x\}) \neq \mathbf{S}^*(\{x\})$.

Now $B \models \mathbf{S}(\{\{x\}\}) = \{\mathbf{S}(\{x\})\} \neq \{\mathbf{S}^*(\{x\})\} = \mathbf{S}^*(\{\{x\}\})$, so if $x$ is a counterexample, so is the object that the model $B$ calls $\{x\}$ – which is the object that $M$ calls $\{j(x)\}$, which has ordinal rank $j(\beta)+1$. It follows that $\beta \leq j(\beta)+1$. In fact, since $\beta = j(\beta)+1$ is impossible (consider the parity of the finite part of the ordinal $\beta$), it follows that $\beta < j(\beta)+1$, so $\beta \leq j(\beta)$.

Since $\beta \leq j(\beta)$ and $\beta < \alpha$, it is clear that $\beta \leq j(\beta) < j(\alpha)$, from which it follows that a counterexample must be a subset of $V_{j(\alpha)}$ so an element of $V_{j(\alpha)+1}$, so a $B$-set.

If $x$ is a counterexample of minimal ordinal rank $\beta$, and $B \models y \in x$, then $M \models j(y) \in x$, so the ordinal rank of $j(y)$ is less than the ordinal rank of $x$: if the ordinal rank of $y$ is $\gamma$, we have $j(\gamma) < \beta$, so we have $\gamma < j^{-1}(\beta) \leq \beta$. It follows that for any $y$ such that $B \models y \in x$, we have $B \models \mathbf{S}(\{y\}) = \mathbf{S}^*(\{y\})$.

Now since $x$ is a $B$-set, we have $B \models \mathbf{S}(\{x\}) = \{\mathbf{S}(\{y\} \mid y \in x\} = \{\mathbf{S}^*(\{y\} \mid y \in x\} = \mathbf{S}^*(\{x\})$, which contradicts our initial assumptions about $x$. We conclude that there can be only one function witnessing the Axiom of Endomorphism in a Boffa model, from which it follows that the membership relation of the original model is definable in the Boffa model, which is what we set out to prove.

## 5 Further Remarks and Conclusions

It is worth noting that, while the notion of ordinal rank used in this argument is an $M$-notion, it is definable in $B$, and in fact we can formulate an assumption about $B$ which is equivalent to the assertion that the universe of $B$ is a rank of the cumulative hierarchy according to $M$. In any well-founded relation $W$, we can define in *NFU* as in ordinary set theory a notion of ordinal rank of elements of the domain of $W$. Further, we can say that an ordinal rank $\beta$ in a relation $W$ is *complete* if every subset of the collection of ordinals of rank $\leq \beta$ is the $W$-preimage of some element of the domain of $W$ of rank $\leq \beta + 1$. Because of the way $B$ is constructed, an assertion satisfied by $B$ is "$E$ is a well-founded relation and any ordinal rank in $E$ is either complete or the entire domain of $E$". *NFU* with the Axiom of Endomorphism and this additional assertion proves that there is at most one function witnessing the truth of the Axiom of Endomorphism. It is much easier to formulate the argument as we have given it in terms of $M$'s notion of rank, as we avoid technicalities about the way ordinals are defined in *NFU* and peculiar properties of the ordinals of *NFU*.

It is further worth remarking that, while it was already well-known (and is easily seen from the consistency proof for *NFU* given above) that *NFU* + Infinity has precisely the consistency strength of Mac Lane set theory, the argument of this paper shows that there is a slight extension of *NFU* (*NFU* + Endomorphism + "$E$ is a well-founded relation and any ordinal rank in $E$ is either complete or the entire domain of $E$") which has a rather more intimate relationship of mutual interpretatibility with Mac Lane set theory (without Infinity) enhanced with an automorphism of the universe which moves a rank: we have seen above that we can derive a model of the extension of *NFU* from the extension of Mac Lane which defines all the concepts of Mac Lane set theory as restricted to its domain. It can further be noted that from the model of *NFU* we can recover an entire model of Mac Lane, not just a rank: the elements of the interpreted Mac Lane are pairs $\langle x, n \rangle$ where $x$ is a set and $n$ is a natural number, with the intention that $\langle x, n \rangle$ code $j^{-n}(x)$; we do not give the (easy) full details of this development here. The model of Mac Lane that is recovered is in effect truncated at the supremum of the ranks indexed by $j^{-n}(\alpha)$'s; this is of course not an ordinal of $M$, and it is important here that we are working in Mac Lane rather than Zermelo set theory, as unbounded quantifiers would be a problem in the latter context. There is another way to interpret Mac Lane set theory with an automorphism in *NFU* (or more accurately the theory of a rank in a model of Mac Lane moved downward by an automorphism): the objects of this interpretation are isomorphism classes of well-founded extensional relations with a top element. The methods used are derived from [6], and full details can be found in [**?**].

It is quite striking that in the usual models of $NFU$, the apparently featureless urelements are thus seen to be far from featureless. A predicate of an urelement $u$ which distinguishes urelements from one another is easily described: $\emptyset \, E \, u$ is a very simple example (its full form in the language of $NFU$ would be the expansion of "there is a function $\mathbf{S}$ witnessing the truth of the Axiom of Endomorphism such that $\emptyset \in \mathbf{S}(\{u\})$"). But this is not the main point of this result: the main point is that one has stronger interpretability of the ambient Mac Lane set theory in which a model of *NFU* is constructed by the Boffa procedure than one would expect.

# References


[1] Boffa, M. "*ZFJ* and the consistency problem for *NF*", in *Jahrbuch der Kurt Gödel Gesellschaft*, 1988, pp. 102-6.

[2] Crabbé, M., "On the set of atoms", *Logic Journal of the IGPL*, vol. 8 (2000), pp. 751-759.

[3] A. Ehrenfeucht and A. Mostowksi, "Models of axiomatic theories admitting automorphisms", *Fundamenta Mathematicae* 43 (1956) pp. 50-68.

[4] T. E. Forster, *Set theory with a universal set, an exploration of an untyped universe*, 2nd. ed., Oxford logic guides, no. 31, OUP, 1995. pp. 68-70.

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

[6] Hinnion, R. *Sur la théorie des ensembles de Quine.* Ph.D. thesis, ULB Brussels, 1975.

[7] 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.

[8] M. Randall Holmes, *Elementary Set Theory with a Universal Set*, volume 10 of the Cahiers du Centre de logique, Academia, Louvain-la-Neuve (Belgium), 1998. 241 pages. ISBN 2-87209-488-1, see chapter 20 for the discussion of well-founded extensional relation types. A revised version is available on the WWW at `http://math.boisestate.edu/~holmes`.

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

[10] Saunders Mac Lane, *Mathematics, Form and Function*, Springer-Verlag, 1986.

[11] A. R. D. Mathias, "The strength of Mac Lane set theory", *Annals of Pure and Applied Logic*, vol 110 (2001), pp. 107-234

[12] Michael Potter, *Set Theory and its Philosophy*, Oxford University Press, 2004.

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

[14] F. Ramsey. "The foundations of mathematics", *Proc. Lond. Math. Soc. 2*, vol. 25, p. 338-384, 1926.

[15] Dana Scott, "Definitions by abstraction in axiomatic set theory", *Bull. Amer. Math. Soc.*, vol. 61 (1955), p. 442.

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

[17] 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.

[18] Wiener, Norbert. "A simplification of the logic of relations", *Proceedings of the Cambridge Philosophical Society*, 17, pp. 387-390, 1914.

[19] H. Wang. *Logic, Computers, and Sets*, p. 406, Chelsea, 1970.

[20] A. N. Whitehead and B. Russell, *Principia Mathematica*, 3 vols, Cambridge University Press, 1910, 1912, and 1913. The abridgment is generally easier to find: Whitehead, Alfred North, and Russell, Bertrand, *Principia Mathematica (to *56)*, Cambridge University Press, 1967.