Symmetry as a criterion for sethood of a class motivating stratified comprehension

M. Randall Holmes

12/22/2016, 4:30 pm

1 Introduction

This is an essay in axiomatics of theories of sets and classes. In the first five sections, we will not be talking formally about models of these theories (at least we do not expect to): if we do such a thing the reader may assume that the metatheory is something standard, such as ZFC. The symbol $\in$ will stand for the membership relation of the theories of sets and classes (or set theories with stratified comprehension) that we discuss.

The sixth section is of a different character, discussing models of given theories of sets and classes in the context of a metatheory which may be taken to be the usual ZFC. **The sixth section is still sketchy.**

It should be noted that this paper is not directly concerned with the vexed problem of the consistency of NF, though it presents a theory which interprets NF. What the paper addresses is the frequent complaint that NF appears to be motivated by a syntactical trick: we present an account of which classes in a theory with sets and proper classes should be sets which is entirely semantic in character and leads to a theory whose sets satisfy the stratified comprehension scheme of NF (or of NFU if atoms are allowed). Showing that this criterion can indeed be implemented is a further problem.

In the last section, we show further that the theory of sets and classes that we develop is a conservative extension of a set theory, NF or NFU with a single additional stratified axiom.

2 The general theory of sets and classes

We begin by presenting a minimal framework for a theory of sets and classes. This is a first-order unsorted theory with membership, equality, and the empty class as primitive notions.

Definition: An object x is a *class* iff $x = \emptyset \vee (\exists y : y \in x)$. A class is either the empty class or has an element. An object x is an *atom* iff it is not a class. We provide abbreviations $\texttt{class}(x)$ and $\texttt{atom}(x)$ for "x is a class" and "x is an atom", respectively.

Axiom of the empty class: $(\forall y : y \notin \emptyset)$.

Axiom of extensionality:

$$(\forall ab : (\exists c : c \in a) \wedge (\forall x : x \in a \leftrightarrow x \in b) \rightarrow a = b).$$

It follows immediately that classes with the same elements are equal (the axiom formally addresses only nonempty classes, but there is only one empty class).

Definition: An object x is an *element* iff x is an atom or $(\exists y : x \in y)$. We write $\texttt{element}(x)$ for "x is an element". A class which is an element is referred to as a *set*. We write $\texttt{set}(x)$ for "x is a set". A class which is not an element is referred to as a *proper class*.

Definition: A formula is *bounded* iff every quantifier appearing in it is restricted to a class.

Axiom scheme of class comprehension: For each bounded formula ϕ and variable A not appearing in ϕ, we assert the axiom

$$(\exists A : (\forall x : \texttt{element}(x) \rightarrow (x \in A \leftrightarrow \phi))).$$

The formula ϕ may contain free variables (parameters).

Definition: It is evident that for each bounded formula ϕ there is a unique *class* witnessing the relevant instance of the axiom scheme of comprehension. We denote this class by $\{x \in V : \phi\}$. The symbol V denotes $\{x \in V : x = x\}$, the class of all elements, which we may refer to as the universal class. Usual extensions of class builder notation will be made without any particular comment necessarily being made.

The restriction to bounded formulas makes this the *predicative* axiom scheme of class comprehension. Removing this restriction gives the impredicative axiom scheme of class comprehension, which we will mention in passing a couple of times below.

Axiom of pairing: For any elements x, y, $\{z \in V : z = x \vee z = y\}$ is a set.

Definition: For elements x, y, we define $\{x, y\}$, the unordered pair of x and y as $\{z \in V : z = x \vee z = y\}$. We define $\{x\}$ as $\{x, x\}$. We define (x, y), the ordered pair of x and y, as $\{\{x\}, \{x, y\}\}$.

Theorem: For any elements x, y, z, w, $(x, y) = (z, w)$ implies $x = z$ and $y = w$.

Proof: For any elements x, y, it is the case that x is the unique object belonging to all elements of (x, y) and y is the unique object belonging to exactly one element of (x, y). Thus, if $(x, y) = (z, w)$, x is the unique object belonging to all elements of (x, y), so it is the unique object belonging to all elements of (z, w), so it is z, and y is the unique object belonging to exactly one element of (x, y), so it is the unique object belonging to exactly one element of (z, w), so it is w.

The axiom of comprehension allows us to associate a class $\{x \in V : P(x)\}$ with each unary predicate P of elements expressible by a bounded formula. The axiom of pairing further allows us to extend this to binary relations on elements: if $x\,R\,y$ is a binary relation on elements expressible by a bounded formula, it can be supposed implemented by the class

$$\{z \in V : (\exists xy : z = (x, y) \wedge x\,R\,y)\},$$

which we will write $\{(x, y) \in V^2 : x\,R\,y\}$.

So far, we have presented the abstract framework for a theory of sets and classes (allowing atoms). The observant reader will notice that the axioms given so far do not allow us to conclude that there are any sets or atoms at all, and only establish the existence of the empty class $\emptyset$.

3 From sets and classes to ordinary set theory via Limitation of Size

A development of a very standard sort of theory of sets and classes can be obtained by adopting the following axioms (which we do not adopt!):

***Axiom of infinity:** There is a set I such that $\emptyset \in I$ and

$$(\forall x : x \in I \rightarrow \{x\} \in I).$$

(This can be formulated in various ways: we choose here to give the original formulation of Zermelo).

***Axiom of power set:** For any set A, $\{x \in V : x \subseteq A\}$, the power class of A, is a set.

***Axiom of union:** For any set A, $\{x \in V : (\exists y \in V : x \in y \wedge y \in A)\}$, the union class of A, is a set.

***Axiom of limitation of size:** A class A is a proper class iff there is a class bijection from A to V.

With these additional axioms, we have a theory of *sets* equivalent to ZFA with a well-ordering of the universe (which of course implies Choice) [strictly speaking, to get ZFA as usually defined we would need to further assert that the class of atoms is a set; our current formulation allows a proper class of atoms, and we would further need to assume Foundation]. The ability to establish Choice is due to a subtlety: the class of set von Neumann ordinals is provably a proper class, so there is a class bijection from the ordinals to V.

As we have noted, we are not adopting the last four axioms. However, we are contemplating one of them. The axiom of Limitation of Size is a criterion for sethood: a class A is a set iff there is not a bijection from A to V, that is, if A is smaller than the universe. What is needed to give flesh to our abstract theory of sets and classes is a criterion for sethood. The abstract theory of sets with the limitation of size criterion (and the other three new axioms to provide some fairly large sets) unfolds into basically the familiar world of standard set theory (iced with proper classes on top).

4 Basic definitions of set theoretic concepts, culminating in the definition of symmetry

Our aim here is to motivate a different picture of the world of sets by framing a different criterion for sethood. Our target set theory is Quine's New Foundations. We will show that a criterion for sethood based on symmetry yields a theory whose sets satisfy NF, just as the theory of sets and classes described above (basically due to von Neumann) yields a theory whose sets satisfy ZFA + Choice [with a possibly proper class of atoms as noted above, and without Foundation]. We in fact develop a theory of sets and classes whose sets satisfy NFU (we allow atoms). Please note that we do not claim that we know how to construct a model of this theory: this is an exercise in axiomatics.

It is very important to remember that from this point forward we are using the language of set theory, but we are talking about a world in which the only things we know about sets and classes are those expressed in the unstarred axioms above. Once we introduce the axiom of symmetric comprehension below, we will acquire more ability to make claims about what sets there are. The only metatheoretical entities we talk about are variables, formulas, natural numbers, and stratifications, which are functions from variables to natural numbers. It should be noted that stratifications could easily be taken to be finite partial functions, though we do not restrict ourselves explicitly in this way, so in fact all of our metatheory could be carried out in arithmetic.

Definition: Let A, B be classes. We define $A \cap B$ as $\{x \in V : x \in A \wedge x \in B\}$ and $A \cup B$ as $\{x \in V : x \in A \wedge x \in B\}$. We define A^c as $\{x \in V : x \notin A\}$.

Definition: If R is a class of ordered pairs, we define R^{-1} as $\{(y, x) \in V^2 : (x, y) \in R\}$. We define $\texttt{dom}(R)$, the domain of R, as $\{x \in V : (\exists y : (x, y) \in R)\}$. We define $\texttt{rng}(R)$, the range of R, as $\texttt{dom}(R^{-1})$.

Definition: We define a *function* as a class of ordered pairs with the property that for any x, y, z, if $(x, y) \in f$ and $(x, z) \in f$, then $y = z$. We write $f : A \rightarrow B$ (f is a function from A to B) if f is a function, $\texttt{dom}(f) = A$, and $\texttt{rng}(f) \subseteq B$. We say that f is an *injection* from A to B iff f is a function from A to B and f^{-1} is a function. We say that f is a surjection from A to B (or a function from A onto B) if $f : A \rightarrow B$

and $\texttt{rng}(f) = B$. We say that f is a bijection from A to B iff f is an injection and a surjection from A to B, or, equivalently, if $f : A \to B$ and $f^{-1} : B \to A$. (Note that we have now given the definitions to support the statement of the axiom of Limitation of Size above). For any function f and $x \in \texttt{dom}(f)$, we define $f(x)$ as the unique y such that $(x, y) \in f$.

Definition: For any classes A, B, we define $A \times B$ as $\{(a, b) \in V^2 : a \in A \wedge b \in B\}$. We define A^2 as $A \times A$. We define $R\lceil A$, for any set of ordered pairs R, as $R \cap (A \times V)$. We define R“A as $\texttt{rng}(R\lceil A)$.

Definition: We define a *permutation* as a bijection from V to V. For any permutation f, we define $j(f)$ as $\{(x, y) : y = f“x \vee \texttt{atom}(x) \wedge y = x\}$. Note that it is not a priori obvious that $y = f“x$ implies $(x, y) \in j(f)$, for x a set, as it may be the case that x is a set but $f“x$ is not. We define $j^0(f)$ as f, and for each concrete natural number n for which $j^n(f)$ is a permutation, we define $j^{n+1}(f)$ as $j(j^n(f))$. Note that if f is a permutation, $j(f)$ is certainly an injection, but it is not necessarily the case that either its domain or its range is V.

Definition: We say that a permutation is *n-setlike* iff $j^n(f)$ is a permutation. We say that a permutation is *setlike* iff it is n-setlike for every n (notice that “setlike” is a meta-theoretic notion, since n is a concrete natural number here, not an internal object of our theory of sets and classes). We say that a class A is *$(n+1)$-symmetric* iff $j^n(f)“(A) = A$ for every n-setlike permutation f. We say that a class A is *symmetric* iff it is $(n + 1)$-symmetric for some n (again, note that “symmetric” is a meta-theoretic notion).

***Criterion for sethood:** A criterion for sethood which we like but which we do not precisely propose here is “A class A is a set precisely if it is symmetric”. The difficulty with this is that we cannot say it. In another paper, we complicated our framework by adding superclasses (with *impredicative* superclass comprehension, where superclasses all of whose elements were sets might fail to be classes) in such a way that we could frame this criterion, and it turns out that (with a restriction of the criterion to avoid an unexpected paradox) it is possible to obtain a theory which entails NF with a criterion motivated in this way. It is unclear whether the system thus obtained is consistent, and it is clearly stronger than NF. It is also quite baroque. Here we do something much simpler with a similar motivation.

We are well aware that it may not be obvious to the reader why a symmetry-based criterion for sethood might be expected to yield a theory of sets and classes whose sets satisfy NF. We will demonstrate why this is to be expected and why it happens.

5 Stratification, the axiom of symmetric comprehension and its consequences

We now introduce the signature notion of NF and related set theories.

Definition: A formula ϕ in the language of first order logic with equality and membership is *stratified* iff there is a function σ from variables (considered as pieces of text) to natural numbers such that $\sigma(x) = \sigma(y)$ for each subformula $x = y$ of ϕ and $\sigma(x)+1 = \sigma(y)$ for each subformula $x \in y$ of ϕ: we are talking about text here, so we should suppose that variables and subformulas here are enclosed in quotes. We call the map σ a *stratification* of ϕ.

Although our language is untyped, we do tend to refer to $\sigma(x)$ as the (relative) type of x in ϕ.

Now we establish that stratification has a nice relationship to symmetry, for which we need to introduce a new syntactic manipulation.

Definition: If ϕ is a bounded formula with stratification σ and f is a permutation which is $\sigma(x)$-setlike for each variable x in ϕ, we define ϕ^f as the result of replacing each occurrence of each variable x in ϕ with $j^{\sigma(x)}(f)(x)$. Strictly speaking we should write something like $\phi^{f,\sigma}$, but we do not. We could justify this by supposing that σ is the unique stratification whose value at each variable x appearing in ϕ is the smallest value at x of any stratification at ϕ (we leave it to the reader to verify that this actually defines a stratification), but what we say makes sense if σ is in each case simply supposed to be understood from context.

Stratification Lemma: If ϕ has stratification σ and has free variables $a_1, \ldots, a_n$ representing sets (we may assume without loss of generality that there are no bound occurrences of the variables a_i in ϕ) and no other free variables, and f is a permutation which is $\sigma(x)$-setlike for each variable x in ϕ, then $\phi \leftrightarrow \phi^f$ and $\phi^f \leftrightarrow \phi_f$, where ϕ_f is obtained by replacing each free variable a_i in ϕ with $j^{\sigma(a_i)}(f)(a_i)$.

Proof of Stratification Lemma: A formula $x = y$ is equivalent to $j^n(f)(x) = j^n(f)(y)$ if f is n-setlike. A formula $x \in y$ is equivalent to $f(x) \in j(f)(x)$ if f is 1-setlike, and by replacing f with $j^n(f)$, is equivalent to $j^n(f)(x) \in j^{n+1}(f)(y)$ if f is $(n+1)$-setlike. It then follows that each atomic subformula $x \, R \, y$ in ϕ, where R is $=$ or $\in$, is equivalent to $j^{\sigma(x)}(f)(x) \, R \, j^{\sigma(y)}(f)(y)$. This establishes $\phi \leftrightarrow \phi^f$. Further, $(\forall x : \psi(x))$ or $(\exists x : \psi(x))$ is equivalent to $(\forall x : \psi(f(x)))$ or $(\exists x : \psi(f(x)))$ respectively for any formula ψ and permutation f. It follows that all decorations of variables in ϕ^f with applications of permutations other than those attached to free variables can be eliminated without changing the truth value of the formula, establishing $\phi^f \leftrightarrow \phi_f$ and the Lemma.

Symmetry Lemma: Let ϕ is a formula with stratification σ, containing free variables $a_1, \ldots, a_n$ representing sets, and f be a permutation which is $\sigma(u)$-setlike for each u in ϕ. We consider the class $\{x \in V : \phi\}$, which we will write in the form $\{x \in V : \phi(x, a_1, \ldots, a_n)\}$. The statement $x \in \{x \in V : \phi(x, a_1, \ldots, a_n)\}$ is equivalent to $\phi(x, a_1, \ldots, a_n)$ which is equivalent to $\phi(j^{\sigma(x)}(f)(x), j^{\sigma(a_1)}(f)(a_1), \ldots, j^{\sigma(a_n)}(f)(a_n))$. This has the interesting effect that $\{x \in V : \phi\}$ is fixed by elementwise application of any $j^{\sigma(x)}(f)$ which satisfies $j^{\sigma(a_i)}(f)(a_i) = a_i$ for each i. Sets which witness stratified instances of comprehension have symmetry of the kind we are discussing, though with a technical qualification.

Definition: We say that a class A is $(n+1)$-symmetric with support a iff $j^n(f)``A = A$ for each n-setlike permutation f such that $j^n(f)(a) = a$.

To exploit the kind of symmetry given by the Symmetry Lemma to get $(n+1)$-symmetry with support, we need to be able to coagulate the conditions $j^{\sigma(a_i)}(f)(a_i) = a_i$ for each i into a single condition of the form $j^{\sigma(x)}(a) = a$.

Observe that any formula $\phi(a_i)$ is equivalent to $(\exists x : x \in a_i' \wedge \phi(x))$, where the new free variable a_i' is intended to refer to $\{a_i\}$, and observe that a stratification σ of ϕ can be modified to σ', stratifying the new formula, with $\sigma'(x) = \sigma(a_i)$ and $\sigma'(a_i') = \sigma(a_i) + 1$: we can as it were raise the types of parameters as desired.

Our next move is to demonstrate that we can merge two parameters of the same type into a single parameter, which shows that we can merge any concrete finite number of parameters into a single parameter, by raising all types to the same value then merging them repeatedly.

Let a, b be two sets. Let c, d, e, f, g, h be six distinct elements. Let $a \oplus b$ be defined as

$$\{\{c\}, \{c, d\}, \{d, e\}, \{e, f\}, \{f, g\}, \{g, h\}\} \cup \{a' \cup \{c, d, e\} : \mathtt{set}(a') \wedge a' \in a\}$$
$$\cup \{b' \cup \{f, g, h\} : \mathtt{set}(b') \wedge b' \in b\}.$$

We claim that for any $n \geq 2$, if $j^n(F)$ is a permutation and fixes this class, $j^n(F)$ fixes a and $j^n(F)$ fixes b. Clearly $j^{n-2}(F)$ must fix c, d, e, f, g, h, because c is the only element of the given class belonging to a singleton, and the concretely given unordered pairs are the only unordered pairs which are elements of the class. It then follows that $j^n(F)$ fixes $\{a' \cup \{c, d, e\} : \mathtt{set}(a') \wedge a' \in a\}$, so for any $a' \in a$ which is a set, $j^{n-1}(F)(a' \cup \{c, d, e\}) = j^{n-1}(F)(a'' \cup \{c, d, e\})$ for some $a'' \in a$, whence it follows that $j^{n-1}(F)(a') = a''$, whence it follows that $j^n(F)(a) = a$. The result $j^n(F)(b) = b$ is proved in the same way.

It remains only to observe that the definition of $a \oplus b$ is stratified, and when parameters a_i, a_j are replaced with a single parameter a_{ij} representing $a_i \oplus a_j$, the value of the modified stratification at a_{ij} will be the same as the common value of the original stratification at a_i and a_j: this construction does not raise type. It is also worth observing that if a is n-symmetric with support s and b is n-symmetric with support t, then $a \oplus b$ is n-symmetric with support $s \oplus t$.

The existence of six distinct elements required for this construction will follow as soon as we have *one* set (which is not its own sole member) or atom, because we have the axiom of pairing: we can take iterated singletons of this set or atom. The existence of one set which is not its own sole member (the empty set) will follow immediately from the axiom of symmetric comprehension we introduce below.

We now enunciate our sethood criterion.

Symmetric comprehension axiom: A class is a set iff it is 3-symmetric or 3-symmetric with support.

The classes $\emptyset$ and V for example are 3-symmetric without support, which gives us at least a couple of sets. As noted above, in combination with

Pairing this gives us the six distinct elements we need to implement merger of support elements.

Spelling out our criterion, a class A is a set iff either $j^2(f)``A = A$ for every 2-setlike permutation $f$ or there is a set $S$, its support, such that for any 2-setlike permutation $f$, $j^2(f)``A = A$ if $j^2(f)(S) = S$. It is useful to note that a set which is 3-symmetric is also 3-symmetric with support $\emptyset$.

Setlikeness Lemma: For every 2-setlike permutation f, $j(f)$ is also a 2-setlike permutation (and so f is 3-setlike, and this can repeated to show that f is n-setlike for any concrete n). This shows that the notion "setlike", which is impossible to express in the language of our abstract framework, is expressible as "2-setlike" in the presence of the axiom of symmetric comprehension.

Proof: That f is 2-setlike means that for every set A, $j^2(f)(A)$ is a set. We want to show that $j(f)$ is 2-setlike, that is, that $j^3(f)(A) = j^2(f)``A$ is a set for each $A$. $A$ is 3-symmetric with some support $S$ by the symmetric comprehension axiom: we deduce from this that $j^2(f)``A$ is 3-symmetric with support $j^2(f)(S)$. Let g be 2-setlike such that $j^2(g)(j^2(f(S))) = j^2(f)(S)$. It follows that $j^2(g \circ f)(S) = S$, whence $j^2(g \circ f)``A = A$, whence $j^2(g)``(j^2(f)``A) = A$, which is what is to be shown. Application of the same reasoning to f^{-1} shows that $j^3(f)$ is a permutation.

Definition: Let $\iota(x)$ be defined as $\{x\}$; the principal use of this notation is that it can be iterated. Observe that $\{\iota(x) : \phi(x)\}$ is definable as $\{u \in V : u = \{x\} \wedge \phi(x))\}$ and that "$u = \{x\} \wedge \phi(x)$" is stratified if ϕ is (after expansion of $u = \{x\}$ into the language of first-order logic with equality and membership). For any concrete k, we can expand $\{\iota^k(x) : \phi(x)\}$ similarly.

First Support Lemma: Let ϕ is a formula with stratification ϕ, containing free variables $a_1, \ldots, a_n$, and f be a permutation which is $\sigma(u)$-setlike for each u in ϕ. We consider the class $\iota^k``\{x \in V : \phi\}$, for a natural number k which we determine below, which we will write in the form $\{\iota^k(x) \in V : \phi(x, a_1, \ldots, a_n)\}$. By considerations above, $\phi(x, a_1, \ldots, a_n)$ is equivalent to a formula $\phi^*(\iota^k(x), a)$ in which a is the only parameter, and the class $\{u \in V : \phi^*(u, a)\}$ is $(\sigma(x) + k + 1) = (\sigma(a) + 1)$-symmetric with support a by considerations given above (we choose k to make the equation hold). The value $\sigma(a)$ is the maximum of the $\sigma(a_i)$'s for the original formula, $\sigma(x)$, and 2. The number k is chosen so that $\sigma(x) + k = \sigma(a)$.

Second Support Lemma: If A is $(n+4)$-symmetric with support S ($n \geq 0$), and S is 3-symmetric with support T, $\bigcup A$ is $(n+3)$-symmetric with support T.

Proof: Let f be a permutation such that $j^{n+2}(f)(T) = T$. It follows that $j^{n+2}(f)``S = S$, that is $j^{n+3}(f)(S) = S$, so $j^{n+3}(f)``A = A$, whence $j^{n+2}(f)``\bigcup A = \bigcup A$. Notice that the Setlikeness Lemma is used here: we need the $(n+2)$-setlike f to be $(n+3)$-setlike as well.

Support Lemma: For any ϕ a formula with stratification σ, containing free variables $a_1, \ldots, a_n$, and f be a permutation which is $\sigma(u)$-setlike for each u in ϕ, $\{x \in V : \phi\}$ is n-symmetric with support for some n, which is shown by expressing it in the form $\bigcup^k(\iota^k``\{x \in V : \phi^*\}$ for suitable ϕ^* and k and applying the First and Second Support Lemmas. The Support Lemma requires the Symmetric Comprehension Axiom, because application of the Second Support Lemma is driven by the assumption that any set appearing as a support is 3-symmetric with a support. It is worth noting that the equations $\bigcup^k(\iota^k``A) = A$ relied on here require the axiom of pairing (ensuring the existence of all required singleton sets).

Comprehension Lemma: Any $(n+4)$-symmetric class is $(n+3)$-symmetric for $n \geq 0$ (and so by repeated application of the lemma a set).

Proof: Suppose that A is an $(n+4)$-symmetric class, so that there is a set S such that for any 2-setlike f (note that by the Setlikeness Lemma, we no longer *strictly* need to decorate "setlike" with a numeral), we have $j^{n+3}(f)``A = A$ iff $j^{n+3}(f)(S) = S$.

Observe that $j^2(f)$ fixes the set $[\subseteq] = \{\{x, y\} : x \subseteq y\}$ iff $f = j(g)$ for some g. We need to show that if $f = j(g)$ is 2-setlike, so is g: since f is 2-setlike, $j(f)$ is a 2-setlike permutation, and $j(f) = j^2(g)$ being a permutation establishes that g is 2-setlike.

Let T be constructed using the parameter merging construction above, such that if $j^{n+2}(f)$ fixes T it fixes $[\subseteq]$ and fixes S. Let f be such that $j^{n+2}(f)$ fixes T. There is then a permutation g such that $j^{n+2}(f) =$

$j^{n+3}(g)$. It follows that $j^{n+3}(g)$ fixes S, so $j^{n+3}(g)“A = j^{n+2}(f)“A = A$, because A is $(n+4)$-symmetric with support S. This shows that A is $(n+3)$-symmetric with support T.

Stratified Comprehension Theorem: Each stratified instance of class comprehension with set parameters is witnessed by a set.

Proof: By the Support Lemma, such a class is n-symmetric with a support for a concrete n. By the Comprehension Lemma, such a class is 3-symmetric with a support, and so is a set by the Symmetric Comprehension Axiom. Where no support is required, $\emptyset$ can always be used as a support.

It follows that the class theory with the Symmetric Comprehension Axiom satisfies NFU (stratified comprehension with weak extensionality) on the domain of sets, while if we assume strong extensionality (there are no atoms) it will interpret NF.

It is useful to note that the Axiom of Pairing (our other very limited criterion for sethood) coheres with stratified comprehension: an unordered pair $\{a, b\}$ is 3-symmetric or 3-symmetric with support iff a and b are 3-symmetric or 3-symmetric with support, by straightforward application of results given above. It may be the case that the Symmetric Comprehension Axiom implies the Axiom of Pairing in the presence of the other axioms, but the availability of Pairing in the abstract framework makes the proofs much more natural.

The main interest of this for us is that this appears to be an entirely semantically motivated development leading to a theory interpreting NF. The syntactical criterion of stratification plays a role in the proofs (in principle dispensible by working with a finite axiomatization of stratified comprehension, it should be noted, but this would not add intelligibility), but plays no role at all in the axiomatization. The value that this paper delivers is a picture of what a world of sets and classes might look like whose sets satisfy NF, not motivated by syntax.

impredicative class comprehension?: We consider the status of our class comprehension axiom: one might wonder why we have chosen the predicative form of von Neumann-Gödel-Bernays class theory rather than the impredicative form of Morse-Kelley set theory. Strengthening the predicative scheme of class comprehension to the impredicative scheme for the system with Limitation of Size gives Morse-Kelley set theory, a natural strengthening of ZFC (mod the possibility of a proper class of atoms). However, strengthening the predicative scheme of class comprehension to the impredicative scheme for the system with Symmetric Comprehension leads to paradox. The difficulty is that it is then possible to define the class of true well-orderings (linear orders every sub*class* of whose fields have minimal elements), and this class would be a set by Symmetric Comprehension, as it is n-symmetric for a small concrete n (depending on how well-orderings are represented as sets), but it cannot be a set: this would lead to the Burali-Forti paradox.

all sets n-symmetric for some n?: Known and rather sophisticated results about NF establish that we must define our symmetry in terms of all permutations, not only set permutations. In NF and related theories, a set A is said to be "cantorian" iff it is equinumerous with ι"A (the bijection between A and ι"A being a set), and "strongly cantorian" iff the restricted singleton map $\iota\lceil A$ is a set. The classes of cantorian sets and strongly cantorian sets are both n-symmetric for small concrete n with respect to permutations which are sets: but they provably cannot be sets, so it is a consequence of our theory with the Symmetric Comprehension Axiom that there are setlike permutations elementwise application of which moves these classes, and these permutations must be proper classes. Further, a strongly cantorian ordinal which is moved to a noncantorian ordinal by a setlike permutation is provably not n-symmetric without support for any n, which reveals that the models of NF(U) derived from SST(U) are not "Forster term models" (a Forster term model being a model in which every element is a set abstract $\{x : \phi\}$ for ϕ a stratified formula without parameters: in such a model, every set is n-symmetric for a concretely given n which can be read from any formula ϕ defining it).

We give the argument in more detail. Since the class of strongly cantorian sets is not a set, there must be a setlike permutation $j^2(f)$ which moves a strongly cantorian ordinal α to a noncantorian ordinal. We choose to represent well-orderings using the set of their initial segments. Under this assumption, it is straightforward to show that any map $j^2(f)$ (f setlike) has a determinate action on order types: two well-orderings of the same type will be sent by $j^2(f)$ to well-orderings of the same type (though if f is not a set map the order type of the images under

$j^2(f)$ may be different). It follows that α is moved by any $j^{n+2}(f)$, because it contains well-orderings of sets of n-fold singletons for every n. Thus, though α is 3-symmetric with support any specific well-ordering of type α, it is not n-symmetric for any n in the strict sense.

A disproof of choice which doesn't work: Here is a sketch of an argument that the symmetric comprehension axiom implies that Choice fails. In fact, it purports to show that there is no linear order of the universe which is a set. We show that in the system with symmetric comprehension it is the case that the support of a linear order of the universe must be rigid (not moved by any setlike permutation other than the identity) and there cannot be a rigid object under symmetric comprehension, as it would serve as a support for any class!

Suppose that L is the linear order with support S (represented as a set of its initial segments). for any setlike permutation f, if $j^2(f)$ fixes S, then $j^3(f)$ fixes L, which implies that $u\,L\,v$ iff $j(f)(u)\,L\,j(f)(v)$. Now suppose that we have a nontrivial orbit in f, the orbit of an object x. For some finite k (either 2 if the orbit is infinite or the size of the orbit otherwise) we will have all sets $X_i = \{f^{i+kn}(x) : n \in \mathbb{N}\}$ distinct and disjoint. $X_i\,L\,j(f)(X_i) = X_{i+1}$ iff $X_{i+1}\,L\,X_{i+2}$ is provable, and leads to contradiction because the order on the X_i's is cyclic.

This seems to imply that no f other than the identity has $j^2(f)$ fixing S, so S is suitable as a support for any class at all, so the Russell class is a set...boom! So this theory appears to have the strong consequence that the universe cannot be linearly ordered.

However, there is a gap in this argument: it is possible that f may have no finite nontrivial orbits and no infinite nontrivial orbit which is a set. If the orbit is a set, the sets X_i can be picked out using L, so it is really necessary that the nontrivial orbit itself be a proper class (not just the classes X_1 and X_2).

The existence of countable proper classes is no surprise here: that is a consequence of NFU.

It is interesting to note that because our criterion for sethood involves an unbounded quantifier, the Symmetric Comprehension Axiom asserts the equivalence of `set`(x) with a formula not permitted to occur in an instance of predicative class comprehension: but notice that the same is true of the criterion of Limitation of Size (the formula equivalent to $\mathtt{set}(x)$ under Limitation of Size asserts that there is no class bijection with a certain property).

6 A converse result?

Still under construction: some of what I say here is speculative and may eventually be retracted. The material above should be correct (mod the inevitable typo here and there).

The reasoning in this section is of a different character: we are discussing models of set and class theories named, the metatheory being some suitable subsystem of ZFC.

We will call the theory decribed above SSTU, for symmetric set theory with urelements. SSTU with strong extensionality (there are no atoms) we will call SST. The sets of SST(U) satisfy NF(U). They also satisfy an additional assertion which NFU does not appear to prove (nor does NF appear to prove it): each set is 3-symmetric with a support, which is a statement which cannot be expressed in the interpreted NF (because it involves a quantifier over possibly proper class permutations) but does imply that every set is 3-symmetric with support with respect to set permutations. We call the theory NF(U) + "each set is 3-symmetric with support relative to set permutations" NFS(U), and we claim that SST(U) is a conservative extension of NFS(U).

This is based on an adaptation of a theorem of Forster and others, to the effect that the stratified sentences are exactly the sentences whose truth values are unaffected by Rieger-Bernays permutation methods. This theorem was actually the motivation for all the work in this paper: it gave a semantic motivation for the notion of stratification, which I set out to adapt into a semantic motivation for the axiom of stratified comprehension.

The class of sentences we consider is a little more general. We allow a stock of function symbols representing setlike permutations (fixing all atoms). This stock of symbols is closed under the operation j and under composition and inverse. The rule for these in relation to stratification is that any occurrence of a term must have the same setlike permutation (or none) applied to it in all contexts in which it appears (where terms are either variables or a function symbol applied to a term). Terms are then assigned relative type in the usual way, and any term must appear with the same relative type wherever it appears in a membership or equality formula. We also provide constants, which do not have to be assigned types for purposes of stratification.

The notion of invariance that we employ is then that a formula $\phi(x_1, \ldots, x_n, a_1, \ldots a_n)$, in which the a_i's are constants, is invariant iff there are numbers τ_i, σ_j such that for any allowable permutation g such that $j^{\sigma_1}[g](a_1) = a_1, \ldots, j^{\sigma_n}[g](a_n) = a_n$ we have $\phi(x_1, \ldots, x_n, a_1, \ldots a_n) \leftrightarrow \phi(j^{\tau_1}[g]x_1, \ldots, j^{\tau_n}[g]x_n, a_1, \ldots a_n)$. We claim that any such formula is equivalent to a stratified formula (in the general sense indicated above).

We have already shown why any stratified formula is invariant in this sense: allowing variables to be adorned with additional function symbols in the way indicated will not change this, and requiring that functions used must fix any constants present avoids difficulties with constants.

Let (M, ϵ_M) and (N, ϵ_N) be two structures for the language of set theory, satisfying [weak] extensionality. Suppose that they satisfy the same stratified sentences. Transform them into models of type theory (M_i, ϵ^τ_M), (N_i, ϵ^τ_N) for the language of type theory (using all integers i as types, including negative integers), in which each type M_i is just a copy $M \times \{i\}$ ofr M and $(x, i)\epsilon^\tau_M(y, i+1)$ iff $x \epsilon_M y$. The models of type theory will of course still satisfy the same stratified sentences.

Use a back and forth argument to construct models $(M^*_i, \epsilon^\tau_{M^*})$, $(N^*_i, \epsilon^\tau_{N^*})$ which are elementary extensions of (M_i, ϵ^τ_M), (N_i, ϵ^τ_N) respectively and which are isomorphic as models of type theory (the language of the type theory not allowing reference to the relation of having the same first component in M or N). We introduce a symbol for a setlike permutation f from M^* to N^*. Where we have not yet created the image under the isomorphism f of $x \in M$ (or later of $x \in M^*$) we first give a complete description of the theory of $(x, 0)$ (including a complete description of its first component in terms of the original theory of M). We then create a new object (or use an existing one if it is determined already) having the correct theory to be $(f(x), 0)$ (this is possible because the two structures have the same stratified theory) then add the new object $f(x)$ to the model N^* (with a complete description of $f(x)$ as an element of N^*), further adding analogues (x, i), $(f(x), i)$ in every type. Then, for the next element of the N or N^* model, define an image under f^{-1} in the same way. Continue this process back and forth until it terminates in a pair of models of type theory with an external isomorphism. The corresponding untyped structures M^* and N^* for the language of set theory will satsify the same stratified sentences, and be obtainable one from the other by a Rieger-Bernays permutation interpretation using f. This establishes that models with the same stratified theory will in fact satisfy the same invariant sentences and vice versa. For any particular invariant sentence ϕ [suitably decorated], the conjunction of all stratified sentences which must hold in any model satisfying ϕ is logically equivalent to ϕ by completeness and must be equivalent to a finite subconjunction by compactness.

Now to build a model of SST [or SSTU] from a model of NF [or NFU], proceed by listing and processing the definable classes paired with candidate supports: if a definable class is not a set, there will be a model (constructed as above) with the same stratified theory in which there is a setlike permutation moving it and fixing the candidate support. If a definable class is a set, it has support which can be read from its definition, and we do nothing except note its support. If there were no such model, it would follow that the formula defining the class was invariant, so stratified, so defined a set. The new permutation is required to respect the supports of objects already defined. Pass to this model (an elementary extension of the previous model) and add the setlike permutation to the language, adding new formulas representing definable classes and support elements to process later. Continue. This process will terminate (even if it leads to a class model, but I believe it will actually terminate in a set model) in a model which satisfies SST(U). It will have the same NF [or NFU] theory as the original model, but the embedded NF model may be much larger than the original one (and have lots more external subclasses than the original one).

I'm well aware that the model theoretical argument in the last two paragraphs needs to be spelled out fully and tidied up.

The additional axiom adjoined to NFS(U) ensures that the sets of the original NF, which are given at the outset rather than defined by formulas, are symmetric, which allows the process described to start. We start out just with the information that the sets of the original model of NFS(U) are symmetric with respect to set permutations, but the process will not introduce any new class permutations which violate this symmetry.

It is worth noting that a "Forster term model" of NF (a model in which every element is an abstract $\{x : \phi\}$ for some parameter-free formula ϕ) has every set n-symmetric for some n without support, from which it follows by analysis given above that every set is 3-symmetric with suitable set support (obtained by merging a finite collection of iterated singletons of $[\subseteq]$ as parameters). From this it follows that the existence of a Forster term model of NF implies the consistency of SST, though the models of NF obtained from

SST cannot be Forster term models.

We comment that this theory appears to have a motivation in rather old-fashioned terms: the idea is that mathematical objects are understood as being constructed by **abstraction**. Passing from a class to its symmetrization relative to setlike permutations fixing a previously given set looks like abstracting a feature from the structure of the class.