Introduction to New Foundations, with attention to errors of Quine

Randall Holmes

Edmonton Logic Group, Halloween, 2023

I haven't really come to bury Quine but to praise him. He did have the idea. He did make some unfortunate mistakes which are good examples of mistakes others can make working with this theory. I don't think NF(U) is so terribly difficult: but it requires a kind of discipline which is not needed in ordinary set theory.

Prehistory of NF: the simple theory of types

New Foundations is an untyped set theory. Its genealogy goes back to a typed set theory though, and the history of this typed set theory is rather vexed and should be interesting to this audience.

The Belgian school of NF-istes called this theory TST and that is what I call it.

TST has types indexed by the natural numbers. Its primitive predicates are equality and membership, with type rules for atomic sentences summarized by the schemata $x^i = y^i$ and $x^i \in y^{i+1}$. We do not adorn every variable with a superscript, but we do require that each variable x has a type $\mathtt{type}(x)$, a natural number, and that the schemata must be followed in constructing atomic sentences.

The axioms are extensionality,

$$(\forall xy : x = y \leftrightarrow (\forall z : z \in x \leftrightarrow z \in y)),$$

and comprehension, $(\exists A : (\forall x : x \in A \leftrightarrow \phi))$. The witness to an instance of comprehension can be written $\{x : \phi\}$: this term has type one higher than that of x.

These are schemes: all well-formed formulas of these shapes are axioms.

It is very amusing that these axioms look exactly like the axioms of naive set theory. The type discipline prevents all disasters.

The intellectual history of this system is interesting. Popular accounts of "Russell's theory of types" often proceed to describe this system, which is not Russell's theory of types. Russell describes something like this in Principles of Mathematics in 1904 (vaguely) but he doesn't know how to do type theory in this way because he doesn't know how to implement ordered pairs as sets: the actual type theory of Principia is vastly more complicated.

I believe Norbert Wiener probably in effect knew about this theory when he defined the ordered pair as a set [not the way we do it] in 1914, but he did not give a formal description. I used to conjecture that the first description of this theory appeared in the New Foundations paper, and I was not far wrong: it appears to have first been described by Tarski [with utterly different notation] in 1930. I believe arithmetic of order ω, this theory with the Peano axioms added for type 0, may have been described a bit earlier, but arithmetic with order ω is not the same theory and in particular does not have the symmetry which motivates New Foundations.

Systematic ambiguity in TST

TST exhibits the phenomenon of systematic ambiguity noted by Russell in the system of PM in a far stronger form.

Provide a map $(x \to x^+)$ on variables, injective and onto variables of positive type, such that x^+ is one type higher than x.

For any formula ϕ, define ϕ^+ as the result of replacing each variable x in ϕ (free or bound) with x^+.

It is easy to check that ϕ^+ is well-formed. More disturbing is the observation that if ϕ is a theorem, so is ϕ^+ (the converse is not true: it can be made true by using all integers as types instead of just natural numbers).

It is also the case that every object $\{x : \phi\}$ has an exact analogue $\{x^+ : \phi^+\}$ in the next higher type.

A signal example is the case of Frege natural numbers. There is a type 2 set which is the collection of all three element subsets of type 0: this is the Frege natural number 3 of type 2. There is a type 17 set which is the collection of all three element subsets of type 15: this is also a Frege natural number three. There is a similar proliferation in PM.

The world of TST looks like a hall of mirrors.

Quine proposed that since these analogous objects at successive types look intellectually like versions of the same thing, and we can prove the same theorems about them (stronger theorems as we go up in type) we should simply say that they are the same. This gives New Foundations.

The description of the theory

New Foundations is an untyped theory with equality and membership as primitive predicates.

Its axioms are extensionality,

$$(\forall xy : x = y \leftrightarrow (\forall z : z \in x \leftrightarrow z \in y)),$$

a single axiom, and comprehension,

$$(\exists A : (\forall x : x \in A \leftrightarrow \phi)),$$

for "stratified" formulas ϕ. The witness to an instance of comprehension can be written $\{x : \phi\}$: this term has type one higher than that of x.

A formula ϕ is said to be stratified iff there is a function σ from variables to natural numbers (or integers) such that 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 $\sigma('x') + 1 = \sigma('y')$. This is exactly the same thing as saying that with a suitable assignment of types to variables, ϕ becomes a well-founded formula of TST.

The stratification criterion of New Foundations prevents us from committing ourselves to the impossible $\{x : x \notin x\}$, because the defining formula is not stratified. It does however commit us to $\{x : x = x\}$ the universal set. We are not in Kansas any more, Toto...

In general, big sets show up in NF. The Frege natural number 3 is definable, and just one of them: the hall of mirrors is collapsed. The universe is a Boolean algebra, with intersections, unions and true complements. This causes a lot of excitement in some minds.

Cardinal numbers can be defined as equivalence classes under the usual relation of equinumerousness. Ordinal numbers can be defined as equivalence classes of well-orderings under isomorphism. You might think this would lay us open to the paradoxes of Cantor and Burali-Forti. It doesn't.

We will describe how the Cantor paradox is defused, partly for its own sake and partly because it is relevant to Quine's first error (and also to his second); the first is the one with the most profound effect on the subsequent history of this kind of set theory.

It cannot be true that $|A| < |\mathcal{P}(A)|$ in New Foundations, because $V = \mathcal{P}(V)$.

It is also the case that $|A| < |\mathcal{P}(A)|$ is a crazy thing to say in New Foundations, because the two occurrences of A are at different types. It is well-formed but peculiar.

Let ι denote the singleton operation $(x \mapsto \{x\})$. [I am not asserting that we have such a function as a set, and in fact this is provably false as we will see in a moment].

The natural assertion in New Foundations (which is also a theorem of TST) is $|\iota"A| < |\mathcal{P}(A)|$, and the proof follows Cantor exactly. $\iota"A$ is used here as a convenient shorthand for the set of one element subsets of A.

Suppose there is a bijection f from $\iota"A$ to $\mathcal{P}(A)$. Consider the set

$$C = \{a \in A : a \notin f(\{a\})\},$$

Note that the definition of C is stratified: it makes sense in TST, something which I can explore on request if you don't see it. Let $\{c\} = f^{-1}(C)$. $c \in C$ is equivalent to $c \notin f^{-1}(\{c\})$, that is, $c \notin C$. This is to say the least deeply unsatisfactory, and establishes our result.

This has the corollary that $\iota"V$ is not the same size as $\mathcal{P}(V) = V$: the singleton operation cannot be implemented as a set, as it would otherwise witness the equivalence in size of these two sets.

Error the First: extensionality is not a freebie

Quine, being a philosopher, knows that the axiom of extensionality is a dubious assertion. We might believe in sets without believing that everything is a set, so it might be more natural to assume that some objects are sets, sets with the same elements are equal, and objects which are not sets, which we might call atoms, have no elements (so they have the same extension as the empty set).

Mathematicians feel the convenience of the assumption that everything is a set (since mathematical objects can in general be implemented as sets, and the axiom is simpler). Quine knows this. He proposes that it is harmless to assume extensionality, because we can arrange for atoms to be their own singletons.

This would actually be done by a permutation process, if it could be done. An atom x has empty extension. The sequence of objects $x, \{x\}, \{\{x\}\} \ldots$ can be dealt with by the following means: redefine membership for all objects of the form $\iota^n(x)$ where x is an atom, so that x becomes its own sole element and each $\iota^n(x)$ becomes its own sole element instead of having $\iota^{n-1}(x)$ as its sole element.

This procedure is very clever and does work to convert models of Zermelo set theory with atoms (the original Zermelo axiomatization allowed them) to models of Zermelo set theory without atoms. But it doesn't work in New Foundations. It fixes extensionality in a model of NFU, the theory with weak extensionality, but it breaks comprehension: the description of the new extension of each object is unstratified, and the interpretation of comprehension cannot be relied upon to be true.

Further, if this did work by a miracle in some model of NFU, it would be a very special one, because the original collection of atoms would become a collection of singleton sets, and so smaller than the collection of sets by Cantor's theorem. In fact, it would be much smaller than the collection of sets, because it would be a collection of n-fold singletons for any concrete natural number n.

NFU is known to be consistent, and the well-known models described by Jensen all have sets of atoms far larger than the collection of sets. It is a theorem of Boffa that the existence of a model of NFU in which there are no more atoms than sets is equivalent to the consistency of NF.

Error the second: Infinity is a theorem of NF but not in the way Quine describes.

Quine claims that Infinity holds in NF by considering the sequence of objects $\emptyset, \{\emptyset\}, \{\{\emptyset\}\} \ldots$ — note that this cannot be shown to be a set using stratified comprehension in any obvious way.

This is a bone headed error and Quine knew better. This definitely establishes that all models of NF are infinite. It does not establish that NF proves that there is an infinite set. I can't overstate this: this was a **serious mistake which a logician should not have made**.

In fact, all models of NFU are infinite, for the reason Quine gives, **and** it is consistent with NFU that the cardinality of the universe is a (nonstandard) natural number.

NF does prove infinity, but the way it proves infinity is scandalous (and not due to Quine). This result is usually presented as a corollary of Specker's 1953 proof that the axiom of choice is false in NF. If the universe is finite, choice is certainly true, so the universe must be infinite.

In fact, Specker originally found a direct proof of infinity independent of the proof of the negation of AC, but very similar in spirit; consideration of this proof very naturally leads to the disproof of AC.

NFU is consistent with Infinity and Choice as was shown by Jensen in 1969. In general NFU is very similar to ordinary set theory in fundamental ways in which NF is not, in spite of the presence of big sets and other weirdness.

NFU is a fairly convenient foundation for mathematics. NF is not, as the failure of Choice reveals, and the consistency problem for NF does not as yet have an accepted resolution.

I have observed that if Quine had not made his first error, and had proposed NFU, the history of this sort of set theory might be quite different. Jensen would have established consistency, and the Specker result would have translated to the assertion that the axiom of choice implies many atoms, which would not be terribly annoying because Jensen's method of construction naturally produces lots of atoms.

The system of Mathematical Logic

In his book Mathematical Logic, Quine enhanced the system of New Foundations with proper classes. For any formula ϕ, there is a class $\{x : \phi\}$ of all **elements** such that ϕ (impredicative class comprehension). The axiom of set comprehension in the first edition then simply asserts that stratified formulas define sets.

Error the third: ML as originally framed falls to the Burali-Forti paradox

I do not have the first edition of ML handy, but I believe the issue is that the first edition allows bound variables ranging over all classes in stratified formulas. The second edition restricts all variables free and bound in instances of stratified comprehension to the universe of sets (the collection of all elements).

This restriction makes ML extend NF only in means of expression: any model of NF then gives a model of ML if the classes of ML are taken to be all of the subsets of the model of NF in the metatheory. Please notice that any analogy with the relationship between Morse-Kelley and ZFC fails: Morse-Kelley does enable strengthenings of the native axioms of ZFC which mention and quantify over proper classes.

How does NF avoid Burali-Forti?

The ordinals are defined in NF as equivalence classes of well-orderings under the usual similarity relation.

The usual order relation on the ordinals is a well-ordering ($\alpha \leq_{\mathrm{ord}} \beta$ iff any element of β (a well-ordering W) has a sub-well-ordering W' whose domain is an initial segment of the domain of W under the order W such that W' belongs to α).

So the order $\leq_{\mathrm{ord}}$ just described belongs to an ordinal Ω. This should begin to make one very nervous.

Now we try out Burali-Forti...the order type of the ordinals $\leq_{\mathrm{ord}}$ Ω should be Ω, but that means there can't be any more ordinals...and it is quite straightforward to construct an element of $\Omega+1$ (there are plenty of things which are not ordinals, append one to the order $\leq_{\mathrm{ord}}$).

The weak point is that the order type of the ordinals less than an ordinal α is not α.

For any relation R, define R^ι as

$$\{(\{x\}, \{y\}) : x\,R\,y\}.$$

Also notice that a relation is three types higher than the elements of its domain and range (in TST terms). [because we are using the Kuratowski pair: this could be a displacement of one, if we used the type level Quine pair, but we are not discussing that brilliant invention here].

If $W \in \alpha$ is a well-ordering, taken to be of type i objects, W is of type $i + 3$. Each element x of the domain of W can be associated with the order type of the order W_x restricting W to $y <_W x$. But this order type is four types higher than x. We can define (in TST and so in NF(U)) an isomorphism between W^{ι^4} and $\leq_{\mathrm{ord}}$ restricted to order types of initial segments of W. It is convenient to define $T(\alpha)$ as the order type of W^ι for $W \in \alpha$. We have outlined the proof that the order type of the restriction of $\leq_{\mathrm{ord}}$ to ordinals $\leq_{\mathrm{ord}} \alpha$ is $T^4(\alpha)$, and so we have shown by the Burali-Forti argument (defanged) that $T^4(\Omega) < \Omega$ (because certainly $T^4(\Omega) + 1 \leq \Omega$).

Note the uncomfortable fact that $\{T^{4i}\Omega\}$ is seen to be a strictly decreasing sequence of ordinals, and so certainly not a set. Any model of NFU contains proper classes which are externally countable!

The same argument can be carried out in ML in its original version, with well-orderings defined using the condition that any proper sub*class* of the domain of a well-ordering has a minimal element, and the precise argument given here leads to the Burali-Forti paradox, because the stronger kind of ordinal thus defined will certainly be well-ordered by the same relation $\leq_{\mathrm{ord}}$ thus restricted and *will* satisfy the common sense result that the order type of the ordinals less than α is α, because transfinite induction can be carried out on this unstratified condition.

Error the fourth: the natural numbers cannot be defined so as to get infinity and math induction as Quine claims

In ML, Quine defines the natural numbers as the intersection of all inductive *classes* of cardinals, which was allowed by the comprehension axiom of the first edition. This enforces mathematicial induction for unstratified conditions, and it also enforces infinity.

Failure of Infinity in NFU entails that $|V|$ is finite, and of course then doesnt have a successor (or doesn't have an inhabited successor, depending on exact definitions). The natural numbers in ML as defined above can be shown to have the successor operation total, so infinity holds.

Unfortunately, Quine didn't notice that this definition of the natural numbers ceases to work when the comprehension axiom is weakened to avoid the Burali-Forti paradox. The problem is that the intersection of all inductive classes of cardinals is then perfectly well definable, but there is no reason to believe it is a set. It could be, and in some models of NF **must** be (if NF is consistent), a proper class initial segment of the natural numbers.

I have to say that this is seriously careless.

One can fix this error and all the errors it causes in subsequent parts of Mathematical Logic by stipulating by axiom that the set $\mathbb{N}$ which is provided by stratified comprehension (the intersection of all inductive sets) coincides with the intersection of all inductive classes.

This is known to strengthen NFU **enormously**, not on an arithmetic level, but on a set theoretical level. It shows the existence of quite large cardinals well above even the cardinality of the reals [it is known to be consistent with NFU]. The same results would hold in NF, but we don't officially know what the strength of NF is. We can still say with confidence that this strengthens NF as well, because the axiom here proposed for ML proves Rosser's Axiom of Counting, the assertion that every finite set A is the same size as $\iota``A$, which is easily proved by an unstratified induction, and Steven Orey showed that if NF is consistent it is consistent with the negation of the Axiom of Counting.