Logique & Analyse **212** (2010), 483–491

NF AT (NEARLY) 75

THOMAS FORSTER

Abstract
The consistency question for Quine's NF is still open. This is despite consistency having been established for systems which apparently resemble it very closely. The peculiar difficulties attending the consistency problem for NF are discussed.

1. *NF and well-foundedness*

NF was presented to the world in a paper of Quine's [8] "New Foundations for Mathematical Logic" in 1937, and it is from that title that NF takes its name. Next year is the seventy-fifth anniversary of the manuscript, and affords an opportunity to reflect on how much has transpired since. As its title suggests, this paper represents an endeavour to survey past achievements and set out current problems.

The first person to make progress with understanding NF was Ernst Specker, who in a series of papers in the 1950's ([9], [10], and [11]) revealed that NF refutes the axiom of choice, and laid bare the connections between NF and type theory. A further leap forward came with Jensen [5] who showed how NFU — which is NF weakened by allowing the existence of urelemente — is consistent. Sadly no discoveries as dramatic as those have been made in recent decades, and the question of the consistency of NF is now the oldest open problem in set theory.[1] Considering its venerability and its philosophical interest the amount of attention it has attracted is surprisingly small. One reason for this is a widespread mistaken feeling that NF is not a theory of *sets* because the sets it concerns itself are in part of a kind that nowadays most people who call themselves set theorists no longer study. It was not always so. Set theory did not start off as a study of "pure" sets (built up purely from other sets). It was rather a study of sets of preëxisting mathematical objects such as reals and real-valued functions, and — more

[1] Perhaps, since there are those who think that CH is in some sense still open, one should perhaps say the oldest open *consistency* problem in set theory.

speculatively — big sets like the universe or the set of all cardinals or the set of all ordinals.

However the focus of interest among set theorists has shifted over the years away from general sets of this kind to pure sets exclusively, and pure *well-founded* sets at that: the cumulative hierarchy. How did this happen? The answer is that we now have *simulacra* within the cumulative hierarchy for all (or at least most) mathematical objects and we are engaged in a pretence that the simulacra *are* the things they simulate. If you want to reason about the set of all functions $\Re \to \Re$ you can now do this while pretending that real numbers are (pure) sets of a particular kind. The feasibility of this pretence makes it possible further to believe that the sets in the cumulative hierarchy are all the sets one needs. Possibly, indeed, all the sets there are.

The world of sets described by NF is a richer and more complex one that harks back to a time when the extra structure — of sets-of-things-other-than-sets, and sets that violate foundation — was embraced. In recent years set theory with antifoundation axioms has attracted a certain amount of attention, but very little of that attention has spilled over into interest in NF, certainly not enough to kick-start NF studies.

So is NF really a theory of sets? Some of the sets studied in NF are illfounded, and Barwise [1] would have us call illfounded sets *hypersets*. The idea would be that a hyperset is a different kind of thing, something a bit like a set, but distinct from it, rather in the way that multisets and lists are a bit like sets, but are distinct from them.[2] This is an error. Let us first establish that at the very least these illfounded-sets/hypersets exist in whatever sense wellfounded sets exist. Since the existence of Quine atoms (objects $x = \{x\}$) and other such is consistent with the remaining axioms of ZF no-one is going to claim that there can be no illfounded sets, and any inclination to believe that in mathematics existence is freedom from contradiction will tell us that Quine atoms exist in whatever sense wellfounded sets do. The scoundrel's last refuge is that they exist all right, it's just that they are a different kind of object…just as multisets, lists, streams *etc* all exist but are different kinds of object from sets.

Modern mathematics supplies us with various suites of objects: sets, lists, groups, rings, fields, vector spaces and so on. These are arbitrary objects-in-extension, so that one can write "Let $\mathbb{F}$ be a field of characteristic 7" and the reader knows what to expect. There will be a set, and it will have structure of a certain kind, appropriate to it being a field rather than, say, a graph. Multisets and sets are as different as fields and topological spaces are different, in that multisets have multiplicity information (which sets do not have) and

[2] I don't think Barwise can have really believed this; he was far too good a mathematician. And he was a very gifted expositor: I think 'hyperset' is nothing more than an inspired piece of marketing.

additionally multisets and lists differ from each other in that lists additionally have order information. Observe however that in this sense there is no difference between illfounded sets ("hypersets") and wellfounded sets. They have the same structure and they do the same things. All you can ask of a set is that it should tell you what its members are (not what their multiplicities are or what operations they support) and that task for the set is the same whether the set is an ordinary wellfounded set or a so-called hyperset. There is nothing operationally to distinguish them. So the set/hyperset distinction is not in the least like the set/multiset distinction. That latter distinction is a genuine type distinction, whereas the distinction between wellfounded sets and illfounded sets is much more like the difference between countable sets and uncountable sets. By all means restrict your study of sets to wellfounded sets — *de gustibus non est disputandum* after all — just don't pretend that those are all the sets there are.

So NF really is a set theory, and is a theory of all sets, not just wellfounded sets. NF has to stay its hand when it comes to the separation principle (that the intersection of a set and a class is a set) lest — since the universe is a set — all classes be sets. NF only says that a subclass of a set is a set if its defining condition is stratified. Although there is no reason to restrict the separation principle where wellfounded sets are concerned NF nevertheless does so. The effect of this is that NF has very little to say about wellfounded sets: in fact as far as we know NF doesn't even prove the existence of wellfounded sets of infinite rank. The problem with NF is that we haven't so far applied sufficient thought to the question of which axioms need to be added to NF to capture properly the conception of sets towards which it is oriented. A historical parallel might be helpful here. The first attempt to axiomatise the view of sets that we now call the cumulative hierarchy was Zermelo's axiomatisation, which is unsatisfactory in many ways. The universal view nowadays is that adding the axiom scheme of replacement was the right thing to do. It's one of those things — like the channel tunnel — that we should have done years earlier. It may be that there are axioms or axiom schemes that should be added to NF which will enable us to see more clearly the picture of the world of sets that it gives us, and to work more easily in it. The currently unsatisfactory nature of NF is no more an argument against the conception of set to which it appeals than the unsatisfactory nature of the Zermelo axiomatisation of set theory was evidence that the concept of wellfounded set was unsatisfactory. One obvious scheme to add would be full separation for wellfounded sets, and possibly replacement for wellfounded sets. Pleasingly both these principles hold in Church-Oswald models, ("C-O") models ([2], [7]) to which we now turn.

2. *Church-Oswald models*

Church-Oswald models are structures that are made out of models of ZF-like theories, and they are models of theories that say there is a universal set.

In the Church-Oswald technique we start with a model of a theory of wellfounded sets, as it might be ZFC. In any sensible theory of this kind we can prove that there is a class function $k : V \longleftrightarrow V \times \{0, 1\}$. We think of $V \times \{0\}$ and $V \times \{1\}$ as two copies of the universe, and we use $V \times \{0\}$ as the original universe and $V \times \{1\}$ as the collection of complements of our original sets, in the following sense. Our new model has the same elements as the old model, and the new membership relation $\in'$ is defined by

$$x \in' y \longleftrightarrow ((\mathtt{snd}(k(y)) = 0 \ \wedge \ x \in \mathtt{fst}(y)) \\ \vee \ (\mathtt{snd}(k(y)) = 1 \ \wedge \ x \notin \mathtt{fst}(y)))$$

Ordered pairs whose first components are 0 correspond to *low* sets: the collection of things that are members of such an ordered pair in the new sense is a set in the old sense. If the first component is 1 the set is *co-low*: the collection of things that are members of such an ordered pair in the new sense is the complement of a set in the old sense. It turns out that the wellfounded sets of the new model form a copy of the (wellfounded) model we started with.[3] This new model thus satisfies an axiom of complementation: for every set x the collection $V \setminus x$ of things not in x is also a set. It satisfies binary union: $x \cup y$ is a set whenever x and y are, and every set x has a singleton $\{x\}$. The set theory asserting these existence principles (plus the axiom of extensionality of course) is called 'NF_2'. This method is initially quite confusing, in that the new models have the same elements but those elements (which are supposed to be sets, after all) acquire novel contents. However once one gets used to it it's quite clear. Essentially what we have done is add names for complements of old sets. No new object (complement) gets created in more than one way. Can we generalise this method? Yes: here are other objects one can add names for by this method: cardinals and generalised cardinals, relational types, and Church does this. In unpublished work Flash Sheridan showed how to make the (graph of the) singleton function into a set. Emerson Mitchell, a Ph.D. student of Church's, showed in [6] how to obtain a model closed under the power set operation. The system NFO has the axioms of NF plus the existence of principal ultrafilters: $B(x) = \{y : x \in y\}$ is a set for all x. (NFO is the subset of NF containing only those set existence axioms where Φ is quantifier-free; the 'O' stands for

[3] At least if certain trivial technical conditions are met.

'Open'. The new existence axiom $(\forall x)(\exists y)(\forall z)(y \in z \longleftrightarrow x \in y)$ is an axiom of the form $(\forall x)(\exists y)(\forall z)(y \in z \longleftrightarrow \Phi)$ where Φ is quantifier-free.) One can also consider stronger systems like $NF\forall$ where Φ may be of the form $(\forall w)\Psi$ where Ψ is quantifier-free. There are C-O models for NFO and NF∀ and other systems as well, but not — as yet — any C-O models for NF. More on this later. For the moment we record two *caveats*.

First caveat: what's new?

The C-O model constructed from a model $\mathfrak{M}$ of ZFC can be seen — by the jaundiced eye — as merely the original model $\mathfrak{M}$ in disguise. The universal set, for example, is really just the empty set with a party hat on. The principal ultrafilter $\{y : \emptyset \in y\}$ is the empty set with a different party hat on. To the jaundiced eye the talk of large sets — the universal set and so on — can therefore be dismissed as mere syntactic sugar for talk about ordinary customary wellfounded sets.

Second caveat: it'll never work anyway

Worse still, if we are to run the C-O construction for a theory T then T must must have a solvable word problem. That is to say, if T says that the universe of sets is closed under certain operations on sets, then any object that can be generated in more than one way by those operations will have more than one name. If the word problem for T is decidable then we can choose one name and discard the others; if we can't — and this will be the case with any sufficiently strong subsystem of NF — there will be chaos. Emerson Mitchell ([6]) spotted this a long time ago:

> "...both the proof in Church's paper and that in this involve constructing exactly one name for each set in the new model. It is easy to construct classes of names for objects satisfying more powerful axioms [...] but, in general, extensionality forces one to set various names equal to each other, which makes other names have the same "members" *et cetera*. Since this kind of model is the opposite of wellfounded there is great difficulty proving that this process converges."

For example, suppose we want the new model to be closed under power-set. We have to create names for power-sets of everything under the sun. But then we find that the names we have for $\{\emptyset\}$ and $\mathcal{P}(\emptyset)$ turn out to name the same thing. This means that we cannot decide what members a (named) set is to have merely by looking at its name — since we do not know which name to use. So we have to be able to detect when things have multiple names so that we can then discard all but one of them. In the C-O model devised by Mitchell every set (not just every low set) had a power set. Observe, however, that $\bigcup$ is left-inverse to $\mathcal{P}$ so if we add a $\bigcup$ constructor then once everything has a power set everything else will acquire two names even if power sets have only one. In fact Mitchell's model is not closed under $\bigcup$, and this is unsatisfactory because $\bigcup$ is such a simple operation.

If nothing ever has more than one name (as in the first generation of C-O models) then of course this problem doesn't arise. If things can have more than one name then we are still all right as long as we can decide when two names name the same thing, for then we can safely retain one single name for an object while discarding all the others — since we know which names they are. We say the theory has *solvable word problem*. The word problem occasioned by adding the power set operation turns out to be solvable, for example.

But if the word problem is unsolvable then we are stuck. Stuck not in the sense that the construction doesn't go through, but stuck in the sense that we have no control and don't know what it constructs.

3. *Will Church-Oswald models ever give a consistency proof for NF?*

It is looking increasingly unlikely. NF can be axiomatised by extensionality plus finitely many axioms saying that the universe of sets is closed under certain simple operations, so on the face of it it is a candidate for a C-O construction. It's true that the word problem for those operations is not obviously solvable, but some ingenuity such as that displayed by Mitchell could in principle come to the rescue. However the problem is deeper than that.

C-O constructions can be used to provide relational types (cardinals, ordinals etc) for low sets. (Church's original model had equipollence classes — cardinals — for all low sets). However C-O constructions never seem to be able to deliver Church-numeral/equivalence-classes/relational types for *all* sets but only for *low* sets, whereas NF needs the set of cardinals of *all* sets and the set of ordinals of *all* wellorderings, and so on. We can add (for example) cardinals for low sets by simply having a third flag (for the cardinal, as well as a flag for the complement). In the same way it's easy enough to add ordinals for wellorderings of low sets, but NO itself (the collection of all, yes *all* ordinals) cannot be low. But adding new flags works only for low sets. In particular getting NO to be a set is a huge problem. The claims NF makes about non-small sets are simply too numerous for comfort. Specker's [9] celebrated refutation of the Axiom of Choice for NF lives entirely on the non-small sets.

There is a further obstacle to obtaining models of NF. Observe that the C-O models of strengthened versions of NF_2 like those we saw above have a kind of recursive structure that is very like the recursive structure enjoyed by the cumulative hierarchy: there is a wellfounded relation that spans the whole of the universe, and everything in the universe can be seen as being defined by recursion over this relation. One might call it the *engendering relation* of the model. In the standard case (ZFC and its congenors) the engendering relation is of course $\in$ — set membership — itself. There is a temptation to think that because it is *sets* that we are trying to study then the engendering relation should be set membership, but there is actually no need for this at all. In the basic Church-Oswald model above the engendering relation $R(x, y)$ is defined by

Either y is low and $x \in y$, or y is co-low and $x \notin y$ (R)

This relation R is wellfounded in the model described above, and many of the purposes served in the standard setting (even forcing!) by the wellfoundedness of $\in$ are served equally well by the wellfoundedness of R instead. What is going on is that the wellfounded relation arises from $\in$ in the original model by considering the operations used to build the words in the theory T. R clearly arises by convolving somehow the $\in$ relation with the complement operation, the characteristic operation of NF_2. If we add the operation B above ("principal ultrafilter") the relation we get is more complicated but it is still a wellfounded one.

In contrast the situation with NF is that there is no even remotely plausible candidate for a definable relation that could engender the universe: the stuff we are looking to the C-O construction to add for us has no recursive structure of the kind the C-O construction relies on.

Does this mean, as Richard Kaye has suggested, that no C-O construction will ever give a model of NF? This thought prompts some interesting reflections. There are two widely-held beliefs:

(i) Set theory is an adequate foundation for mathematics
(ii) Set theory is the study of the wellfounded sets.

The conjunction of these two is what one might call the *mainstream foundationalist* view. (The point is not that foundationalism is mainstream, but that this is the mainstream view among foundationalists. It's syncategorematic rather than attributive). This conjunction has the consequence that everything mathematical that can be constructed can be constructed as a wellfounded set by the methods used to construct wellfounded sets, namely transfinite recursion.[4]

So far so good. Suppose now that NF is consistent. (It might be, for all we know). Suppose also that the mainstream foundationalists are right. What could a construction of a model for NF conceivably look like if they are? It would have to be a transfinite construction of some kind, executed inside the cumulative hierarchy, along the lines of the construction in [3] only much more sophisticated. But such an engine can be nothing but a C-O construction. So if Kaye is correct in his hunch that there is no C-O model of NF then there can be no model of NF at all.

4. *Conclusion*

In this brief essay I have concentrated mainly on the consistency question for NF, and have said nothing about how the axioms of NF can be motivated. An explanation of the roots of stratification — and an explanation of why it is not a mere *ad hoc* syntactic trick — requires much a more extended treatment. An extended treatment can be given, too, of NFU, since this is known to be consistent. The relation between models of NFU and nonstandard models of fragments of ZF is a fascinating area which is not sufficiently widely appreciated. Sadly that, too, is too technical for a treatment here.

Department of Pure Mathematics and Mathematical Statistics
Centre for Mathematical Sciences
Wilberforce Road
Cambridge
CB3 0WB
United Kingdom
E-mail: T.Forster@dpmms.cam.ac.uk

REFERENCES

[1] Jon Barwise and Laurence Moss. *Vicious Circles*. Cambridge University Press. 1996. ISBN.1-57586-009-0 (paper) 1-57586-008-2 (pbk) 202 pp.

[4] Note that this position does not deny the existence of illfounded sets or other objects of similar flavour such as are to be found in [1]. After all, the C-O constructions give us sets like the universal set and equivalence classes of low sets under equipollence. But in these cases there is the wellfounded engendering relation lurking in the background that is doing all the work.

[2] Church, A. Set theory with a Universal Set. *Proceedings of the Tarski Symposium.* Proceedings of Symposia in Pure Mathematics XXV, ed. L. Henkin, Providence, RI, pp. 297–308. Also in *International Logic Review* 15, pp. 11–23.

[3] Forster, T.E. The Iterative Conception of Set. *Review of Symbolic Logic* 1, pp. 97–110.

[4] Forster, T.E. Church-Oswald models for Set Theory. In *Logic, Meaning and Computation: essays in memory of Alonzo Church*, Synthese library 305. Kluwer, Dordrecht, Boston and London 2001. There is a more up-to-date version on www.dpmms.cam.ac.uk/~tf/church2001.pdf.

[5] Jensen, R.B. On the consistency of a slight(?) modification of Quine's NF. *Synthese* 19, pp. 250–63.

[6] Emerson Mitchell. A model of Set theory with a Universal set. Ph.D. thesis, University of Winsconsin 1976.

[7] Oswald, U. Fragmente von "New Foundations" und Typentheorie. Ph.D. thesis, ETH Zürich.

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

[9] 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* 39, pp. 972–5.

[10] Specker, E.P. Dualität. *Dialectica* 12, pp. 451–465. (English translation on www.dpmms.cam.ac.uk/~tf/dualityquinevolume.pdf.)

[11] Specker, E.P. Typical ambiguity. In *Logic, methodology and philosophy of science*. Ed. E. Nagel, Stanford.