New Foundations is consistent

M. Randall Holmes and Sky Wilshaw

April 23, 2024

Contents

1 Introduction

1.1 Introductory note

We are presenting an argument for the consistency of Quine's set theory New Foundations (NF) (see [12]). The consistency of this theory relative to the usual systems of set theory has been an open question since the theory was proposed in 1937, with added urgency after Specker showed in 1954 that NF disproves the Axiom of Choice (see [18]). Jensen showed in 1969 that NFU (New Foundations with extensionality weakened to allow urelements) is consistent and in fact consistent with Choice (and also with Infinity and with further strong axioms of infinity) (see [10]).

The first author showed the equiconsistency of NF with the quite bizarre system TTT (tangled type theory) in 1995 (see [6]; I really don't recommend this paper; the presentation of its results here is better), which gave a possible approach to a consistency proof. Since 2010, the first author has been attempting to write out proofs, first of the existence of a tangled web of cardinals in Mac Lane set theory [an indication of this approach is given below] and then directly of the existence of a model of tangled type theory. These proofs are difficult to read, insanely involved, and involve the sort of elaborate bookkeeping which makes it easy to introduce errors every time a new draft is prepared. The second author (with the assistance of others initially) has formally verified the proof of the first author of the existence of a model of TTT (see [23]), and so of the consistency of New Foundations, in the Lean proof verification system, and in the process has suggested minor corrections and considerable formal improvements to the argument originally proposed by the first author. The second author reports that the formalized proof is still difficult to read and insanely involved with nasty bookkeeping. Both authors feel that there ought to be a simpler approach, but the existing argument at least strongly resists attempts at simplification.

All the theories mentioned here are discussed (and referenced) in the next section.

Any remarks in the first person singular may be attributed to the first author.

1.2 Acknowledgements

1.3 Dated remarks on versions

initial note for this series: The formal verification of the results in section 3 and 4 is complete. Sky's Lean formalization concludes with a proof that there is a structure for the language of TTT which satisfies all typed versions of the statements in the Hailperin axiomatization of NF. Such a structure is a model of TTT, and the existence of a model of TTT implies the consistency of New Foundations. Thus we know that the construction in the paper is correct in general terms and we need to rewrite it for publication to parallel the formalized proof as much as possible: we certainly know that the conclusion of the paper expressed in the title is true.

   initial note for the previous series: This document is probably my best overall version so far. The immediate occasion for its preparation was to serve students attempting to verify the proof in Lean. A formal verification should avoid metamathematics, so it is the fact that the structure defined in section 3 is a model of TTT which should be verified, and further, a finite axiomatization (mod type indexing) of TST and thus TTT should be verified in the model in lieu of the usual statement of the axiom of comprehension of TTT. **This program has been completed, on the Lean side.**

4/23/2024: 11 pm: deleted a redundant paragraph in the counting argument.

   3:37 pm completed bibliography entries. Added a statement about Sky's contribution to the definition of the structure in proposing use of supported model elements rather than simply model elements as domain elements of the f maps.

   10:33 am: edited citations and added all bibliography entries except ones to software.

   9:35 am: noticed and fixed a really stupid mistake: the crucial inequality in the inductive hypothesis in FoA had accidentally been edited out!

   as of 9:09 Boise time, further tweaks to the Hailperin material and a more explicit statement at the end of section 4.4 about how the axiom of type lowering is proved in the formalization.

   as of 8:39 am Boise time, cleaned up the material on the Hailperin axiomatization a bit.

4/22/2024: as of 12:55 am the next day, definition of quantification in the Hailperin treatment was incorrect, is now incomplete. The proof doesn't hinge on this at all (the Hailperin axiomatization is known to work), but I'd like to get it straight. 12:57 maybe I got it. Going to sleep, tomorrow is another day.

   as of 11:36 pm, corrected oversimplifications of the discussion of why the Hailperin axioms work. They are quite horrible and there are probably still indexing errors in my development.

   as of 4:30 pm, added a final section to section 2 on the Hailperin axiomatization in the version actually used in the formalization and comments

on its use in the formal proof in the last subsection of section 4.

The new material on the Hailperin axioms is quite dense and needs reading.

4/21/2024: as of 1147 am, posted revised text for the Freedom of Action Theorem proof, emphasizing the point that a recursion on ι_* drives the order in which computations are made in every case.

as of 9 am Boise time, Minor repairs to the text. The definition of ι_*^+ was stated incorrectly, but easy to fix as the correct definition is everywhere presumed. Other things more minor.

4/20/2024: 4 PM Boise time: removed comments from before the completion of the formalization and reordered them so they can be added at the top from now on.

4/20/24: FoA proof looks pretty clean. Added statement explaining the overall structure of the recursion in terms of the position function.

2:37 Boise time: finished a reading pass through the rest of the paper with minor revisions which made me reasonably happy. I'm sure there are still bugs, and I'm waiting for Sky to read through it.

I still endorse the following statement but I reduced some of the mess:

The paper proof is still a mess. There are definitely conceptual slips (which by courtesy may be called typos) in instances of complex notations in the text as it stands, and further there are difficulties caused by multiple bouts of revisions of different parts of the text, lately driven by the demands of the formalization. Cleaning it up remains a daunting task. However, the formalized proof, which was built by reading an earlier version of the paper proof and coding it into Lean, proves the desired result. This means that cleanup will converge, if pursued diligently, to a correct text.

4/19/2024: Eliminated redundant components from specifications of strong supports.

As of 6:54 am, the part of the counting argument after the lemma correlating specifications with orbits in the strong supports is NOT written correctly, as it is not adjusted to reflect the return to use of strong supports in this version. There isn't an essential difficulty (it is so nice to know the formalization is there...) but some rewriting is needed, which will probably be accomplished today.

As of 9:33 am, support definition restored to what it was and proof of Freedom of Action substantially corrected. The change in the definition of approximation required a substantial addition to the proof and there were serious typos.

As of 12:42 pm I believe sanity is restored to the text of the counting argument (fingers crossed).

As of 7:24 pm, reviewed section 3 which looks very clean. Fixed the description of the recursion in the proof of the inflexible case in Freedom of Action.

4/18/2024: 1240 am Boise time, finished pass through material already worked on, up to the end of the lemma correlating specifications with orbits in the strong supports.

An essential condition in the definition of approximations was left out in editing: fixed

I observe that I seem to be using the condition that nearby litters in the range of a support have all atoms in their symmetric difference also in the range of the support in the proof of Freedom of Action, so I moved that clause to the definition of support proper. I think this is harmless, but it may occasion discussion.

12:58 Boise time: very few changes made in the entire rest of the paper – I made some edits but nothing much is affected by the major changes in definitions made earlier. I'll be continuing to read your reconciliation document to see if there is more I need to address in these sections, and I don't imagine that the editing process is finished at all!

.

4/17/2024: FoA proof looks reasonably clean. Added a note giving an overall explanation of the structure of the recursion in terms of position function. My old remarks on the structure of the recursion were so involved as to be incomprehensible and I think actually incorrect.

1:03 PM Boise time, revisions to section 3 which I believe eliminate all notes from the reconciliation process, and correction of some errors in the text of the proof of Freedom of Action.

5 PM Boise time: revisions complete to the end of the Lemma correlating specifications with orbits in strong supports.

4/16/2024: Added Sky as an author and added a statement about the history of the paper and the roles of the authors.

4/15/2024: This note is attached to a non public version of this paper which has been edited up to just after the definition of specification in section 4 to coordinate with the formalization. Things are looking good.

4/14/2024: Working systematically through a document Sky prepared for reconciliation of the formalization and the paper. Fixed things up to the end of the proof of uniqueness of distinguished extensions, while making systematic changes in the machinery driven by the Lean formalization.

4/11/2024: The formal verification is complete. I am working on editing this paper to be more or less in line with the Lean formal proof (correcting slips and vaguenesses in this document found during the verification).

specific notes for this edit: the result about the cardinality of the collection of near-litters is correct but Sky says needs slightly more attention for proof.

Sky reintroduced strong supports, and there is some finickiness about exactly how they are defined.

There is a slip in the phrasing of the last part (axiom of union); standard stuff about how inflexible litters are handled, I just skated over it.

All is well, because it all works in Lean.

2 Development of relevant theories

2.1 The simple theory of types TST and TSTU

We introduce a theory which we call the simple typed theory of sets or TST, a name favored by the school of Belgian logicians who studied NF (*théorie simple des types*). This is not the same as the simple type theory of Ramsey and it is most certainly not Russell's type theory (see historical remarks below).

TST is a first order multi-sorted theory with sorts (types) indexed by the nonnegative integers. The primitive predicates of TST are equality and membership.

The type of a variable x is written $\mathtt{type}(\text{`}x\text{'})$: this will be a nonnegative integer. A countably infinite supply of variables of each type is supposed. An atomic equality sentence '$x = y$' is well-formed iff $\mathtt{type}(\text{`}x\text{'}) = \mathtt{type}(\text{`}y\text{'})$. An atomic membership sentence '$x \in y$' is well-formed iff $\mathtt{type}\text{`}(x\text{'}) + 1 = \mathtt{type}(\text{`}y\text{'})$.

The axioms of TST are extensionality axioms and comprehension axioms.

The extensionality axioms are all the well-formed assertions of the shape $(\forall xy : x = y \leftrightarrow (\forall z : z \in x \leftrightarrow z \in y))$. For this to be well typed, the variables x and y must be of the same type, one type higher than the type of z.

The comprehension axioms are all the well-formed assertions of the shape $(\exists A : (\forall x : x \in A \leftrightarrow \phi))$, where ϕ is any formula in which A does not occur free.

The witness to $(\exists A : (\forall x : x \in A \leftrightarrow \phi))$ is unique by extensionality, and we introduce the notation $\{x : \phi\}$ for this object. Of course, $\{x : \phi\}$ is to be assigned type one higher than that of x; in general, term constructions will have types as variables do.

The modification which gives TSTU (the simple type theory of sets with urelements) replaces the extensionality axioms with the formulas of the shape

$$(\forall xyw : w \in x \rightarrow (x = y \leftrightarrow (\forall z : z \in x \leftrightarrow z \in y))),$$

allowing many objects with no elements (called atoms or urelements) in each positive type. A technically useful refinement adds a constant $\emptyset^i$ of each positive type i with no elements: we can then address the problem that $\{x^i : \phi\}$ is not uniquely defined when ϕ is uniformly false by defining $\{x^i : \phi\}$ as $\emptyset^{i+1}$ in this case.

2.1.1 Typical ambiguity

TST(U) exhibits a symmetry which is important in the sequel.

Provide a bijection $(x \mapsto x^+)$ from variables to variables of positive type satisfying $\mathtt{type}(x^+) = \mathtt{type}(x) + 1$.

If ϕ is a formula, define ϕ^+ as the result of replacing every variable x (free and bound) in ϕ with x^+ (and occurrences of $\emptyset^i$ with $\emptyset^{i+1}$ if this is in use). It should be evident that if ϕ is well-formed, so is ϕ^+, and that if ϕ is a theorem, so is ϕ^+ (the converse is not the case). Further, if we define a mathematical object as a set abstract $\{x : \phi\}$ we have an analogous object $\{x^+ : \phi^+\}$ of the next higher type (this process can be iterated).

7

The axiom scheme asserting $\phi \leftrightarrow \phi^+$ for each closed formula ϕ is called the Ambiguity Scheme. Notice that this is a stronger assertion than is warranted by the symmetry of proofs described above.

2.1.2 Historical remarks

TST is not the type theory of the *Principia Mathematica* of Russell and Whitehead ([15]), though a description of TST is a common careless description of Russell's theory of types.

Russell described something like TST informally in his 1904 *Principles of Mathematics* ([14]). The obstruction to giving such an account in *Principia Mathematica* was that Russell and Whitehead did not know how to describe ordered pairs as sets. As a result, the system of *Principia Mathematica* has an elaborate system of complex types inhabited by n-ary relations with arguments of specified previously defined types, further complicated by predicativity restrictions (which are in effect cancelled by an axiom of reducibility). The simple theory of types of Ramsey eliminates the predicativity restrictions and the axiom of reducibility, but is still a theory with complex types inhabited by n-ary relations.

Russell noticed a phenomenon like the typical ambiguity of TST in the more complex system of *Principia Mathematica*, which he refers to as "systematic ambiguity".

In 1914 ([22]), Norbert Wiener gave a definition of the ordered pair as a set (not the one now in use) and seems to have recognized that the type theory of *Principia Mathematica* could be simplified to something like TST, but he did not give a formal description. The theory we call TST was apparently first described by Tarski in 1930 ([20]).

It is worth observing that the axioms of TST look exactly like those of "naive set theory", the restriction preventing paradox being embodied in the restriction of the language by the type system. For example, the Russell paradox is averted because one cannot have $\{x : x \notin x\}$ because $x \in x$ (and so its negation $\neg x \in x$) cannot be a well-formed formula.

It was shown around 1950 (in [9]) that Zermelo set theory proves the consistency of TST with the axiom of infinity; TST + Infinity has the same consistency strength as Zermelo set theory with separation restricted to bounded formulas.

2.2 Some mathematics in TST; the theories TST_n and their natural models

We briefly discuss some mathematics in TST.

We indicate how to define the natural numbers. We use the definition of Frege (n is the set of all sets with n elements). 0 is $\{\emptyset\}$ (notice that we get a natural number 0 in each type $i+2$; we will be deliberately ambiguous in this discussion, but we are aware that anything we define is actually not unique, but reduplicated in each type above the lowest one in which it can be defined). For any set A at all we define $\sigma(A)$ as $\{a \cup \{x\} : a \in A \wedge x \notin a\}$. This is definable for any A of type $i+2$ (a being of type $i+1$ and x of type i). Define 1 as $\sigma(0)$, 2 as $\sigma(1)$, 3 as $\sigma(2)$, and so forth. Clearly we have successfully defined 3 as the set of all sets with three elements, without circularity. But further, we can define $\mathbb{N}$ as $\{n : (\forall I : 0 \in I \wedge (\forall x \in I : \sigma(x) \in I) \rightarrow n \in I\}$, that is, as the intersection of all inductive sets. $\mathbb{N}$ is again a typically ambiguous notation: there is an object defined in this way in each type $i+3$.

The collection of all finite sets can be defined as $\bigcup \mathbb{N}$. The axiom of infinity can be stated as $V \notin \bigcup \mathbb{N}$ (where $V = \{x : x = x\}$ is the typically ambiguous symbol for the type $i+1$ set of all type i objects). It is straightforward to show that the natural numbers in each type of a model of TST with Infinity are isomorphic in a way representable in the theory.

Ordered pairs can be defined following Kuratowski and a quite standard theory of functions and relations can be developed. Cardinal and ordinal numbers can be defined as Frege or Russell would have defined them, as isomorphism classes of sets under equinumerousness and isomorphism classes of well-orderings under similarity.

The Kuratowski pair $(x, y) = \{\{x\}, \{x, y\}\}$ is of course two types higher than its projections, which must be of the same type. There is an alternative definition (due to Quine in [11]) of an ordered pair $\langle x, y \rangle$ in TST + Infinity which is of the same type as its projections x, y. This is a considerable technical convenience but we will not need to define it here. Note for example that if we use the Kuratowski pair the cartesian product $A \times B$ is two types higher than A, B, so we cannot define $|A| \cdot |B|$ as $|A \times B|$ if we want multiplication of cardinals to be a sensible operation. Let ι be the singleton operation and define $T(|A|)$ as $|\iota``A|$ (this is a very useful operation sending cardinals of a given type to cardinals in the next higher type which seem intuitively to be the same; also, it is clearly injective, so has a (partial) inverse operation T^{-1}). The definition of cardinal multiplication if we use the Kuratowski pair is then $|A| \cdot |B| = T^{-2}(|A \times B|)$. If we use the Quine pair this becomes the usual definition $|A| \cdot |B| = |A \times B|$. Use of the Quine pair simplifies matters in this case, but it should be noted that the T operation remains quite important (for example it provides the internally representable isomorphism between the systems of natural numbers in each sufficiently high type).

Note that the form of Cantor's Theorem in TST is not $|A| < |\mathcal{P}(A)|$, which would be ill-typed, but $|\iota``A| < |\mathcal{P}(A)|$: a set has fewer unit subsets than subsets. The exponential map $\exp(|A|) = 2^{|A|}$ is not defined as $|\mathcal{P}(A)|$, which would

be one type too high, but as $T^{-1}(|\mathcal{P}(A)|)$, the cardinality of a set X such that $|\iota\text{``}X| = |\mathcal{P}(A)|$; notice that this is partial. For example $2^{|V|}$ is not defined (where $V = \{x : x = x\}$, an entire type), because there is no $X$ with $|\iota\text{``}X| = |\mathcal{P}(V)|$, because $|\iota\text{``}V| < |\mathcal{P}(V)| \leq |V|$, and of course there is no set larger than V in its type.

For each natural number n, the theory TST_n is defined as the subtheory of TST with vocabulary restricted to use variables only of types less than n (TST with n types). In ordinary set theory TST and each theory TST_n have natural models, in which type 0 is implemented as a set X and each type i in use is implemented as $\mathcal{P}^i(X)$. It should be clear that each TST_n has natural models in bounded Zermelo set theory, and TST has natural models in a modestly stronger fragment of ZFC.

Further, each TST_n has natural models in TST itself, though some care must be exercised in defining them. Let X be a set. Implement type i for each $i < n$ as $\iota^{(n-1)-i}\text{``}\mathcal{P}^i(X)$. If X is in type j, each of the types of this interpretation of TST_n is a set in the same type $j + n - 1$. For any relation R, define R^{ι} as $\{(\{x\}, \{y\}) : xRy\}$. The membership relation of type $i - 1$ in type i in the interpretation described is the restriction of $\subseteq^{\iota^{(n-1)-i}}$ to the product of the sets implementing type $i - 1$ and type i.

Notice then that, for each concrete natural number n, we can define truth for formulas in these natural models of TST_n in TST, though not in a uniform way which would allow us to define truth for formulas in TST in TST.

Further, both in ordinary set theory and in TST, observe that truth of sentences in natural models of TST_n is completely determined by the cardinality of the set used as type 0. since two natural models of TST or TST_n with base types implemented by sets of the same cardinality are clearly isomorphic.

2.3 New Foundations and NFU

In [12], 1937, Willard van Orman Quine proposed a set theory motivated by the typical ambiguity of TST described above. The paper in which he did this was titled "New foundations for mathematical logic", and the set theory it introduces is called "New Foundations" or NF, after the title of the paper.

Quine's observation is that since any theorem ϕ of TST is accompanied by theorems $\phi^+, \phi^{++}, \phi^{+++}, \ldots$ and every defined object $\{x : \phi\}$ is accompanied by $\{x^+ : \phi^+\}, \{x^{++} : \phi^{++}\}, \{x^{+++} : \phi^{+++}\}$, so the picture of what we can prove and construct in TST looks rather like a hall of mirrors, we might reasonably (?) suppose that the types are all the same.

The concrete implementation follows. NF is the first order unsorted theory with equality and membership as primitive with an axiom of extensionality $(\forall xy : x = y \leftrightarrow (\forall z : z \in x \leftrightarrow z \in y))$ and an axiom of comprehension $(\exists A : (\forall x : x \in A \leftrightarrow \phi))$ for each formula ϕ in which A is not free which can be obtained from a formula of TST by dropping all distinctions of type. We give a precise formalization of this idea: provide a bijective map $(x \mapsto x^*)$ from the countable supply of variables (of all types) of TST onto the countable supply of variables of the language of NF. Where ϕ is a formula of the language of TST, let ϕ^* be the formula obtained by replacing every veriable x, free and bound, in ϕ with x^*. For each formula ϕ of the language of TST in which A is not free in ϕ^* and each variable x^*, an axiom of comprehension of NF asserts $(\exists A : (\forall x^* : x^* \in A \leftrightarrow \phi^*))$.

In the original paper, this is expressed in a way which avoids explicit dependence on the language of another theory. Let ϕ be a formula of the language of NF. A function σ is a stratification of ϕ if it is a (possibly partial) map from variables to non-negative 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 $\sigma('x') + 1 = \sigma('y')$. A formula ϕ is said to be stratified iff there is a stratification of ϕ. Then for each stratified formula ϕ of the language of NF and variable x we have an axiom $(\exists A : (\forall x : x \in A \leftrightarrow \phi))$. The stratified formulas are exactly the formulas ϕ^* up to renaming of variables.

NF has been dismissed as a "syntactical trick" because of the way it is defined. It might go some way toward dispelling this impression to note that the stratified comprehension scheme is equivalent to a finite collection of its instances, so the theory can be presented in a way which makes no reference to types at all. This is a result of Hailperin ([4]), refined by others, and discussed in some detail in a subsequent subsection of this section, because it is relevant to the Lean formalization. One obtains a finite axiomatization of NF by analogy with the method of finitely axiomatizing von Neumann-Gödel-Bernays predicative class theory. It should further be noted that the first thing one does with any finite axiomatization is prove stratified comprehension as a meta-theorem, in practice, but it remains significant that the theory can be axiomatized with no reference to types at all. It is also worth noting that the collection of all typed instances of the Hailperin axioms is an axiomatization of TST, not a finite axiomatization, but an axiomatization with finitely many templates.

For each stratified formula ϕ, there is a unique witness to

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

(uniqueness follows by extensionality) whch we denote by $\{x : \phi\}$.

Jensen in [10], 1969 proposed the theory NFU which replaces the extensionality axiom of NF with

$$(\forall xyw : w \in x \rightarrow (x = y \leftrightarrow (\forall z : z \in x \leftrightarrow z \in y))),$$

allowing many atoms or urelements. One can reasonably add an elementless constant $\emptyset$, and define $\{x : \phi\}$ as $\emptyset$ when ϕ is false for all x.

Jensen showed that NFU is consistent and moreoever NFU + Infinity + Choice is consistent. We will give an argument similar in spirit though not the same in detail for the consistency of NFU in the next section.

An important theorem of Specker ([19], 1962) is that NF is consistent if and only if TST + the Ambiguity Scheme is consistent. His method of proof adapts to show that NFU is consistent if and only if TSTU + the Ambiguity Scheme is consistent. Jensen used this fact in his proof of the consistency of NFU. We indicate a proof of Specker's result using concepts from this paper below.

In [18], 1954, Specker had shown that NF disproves Choice, and so proves Infinity. At this point if not before it was clear that there is a serious issue of showing that NF is consistent relative to some set theory in which we have confidence. There is no evidence that NF is any stronger than TST + Infinity, the lower bound established by Specker's result.

Note that NF or NFU supports the implementation of mathematics in the same style as TST, but with the representations of mathematical concepts losing their ambiguous character. The number 3 really is realized as the unique set of all sets with three elements, for example. The universe is a set and sets make up a Boolean algebra. Cardinal and ordinal numbers can be defined in the manner of Russell and Whitehead.

The apparent vulnerability to the paradox of Cantor is an illusion. Applying Cantor's theorem to the cardinality of the universe in NFU gives $|\iota``V| < |\mathcal{P}(V)| \leq |V|$ (the last inequality would be an equation in NF), from which we conclude that there are fewer singletons of objects than objects in the universe. The operation $(x \mapsto \{x\})$ is not a set function, and there is every reason to expect it not to be, as its definition is unstratified. The resolution of the Burali-Forti paradox is also weird and wonderful in NF(U), but would take us too far afield.

2.4 Tangled type theory TTT and TTTU

In [6], 1995, this author described a reduction of the NF consistency problem to consistency of a typed theory, motivated by reverse engineering from Jensen's method of proving the consistency of NFU.

Let λ be a limit ordinal. It can be ω but it does not have to be.

In the theory TTT (tangled type theory) which we develop, each variable x is supplied with a type $\mathtt{type}('x') < \lambda$; we are provided with countably many distinct variables of each type.

For any formula ϕ of the language of TST and any strictly increasing sequence s in λ, let ϕ^s be the formula obtained by replacing each variable of type i with a variable of type $s(i)$. To make this work rigorously, we suppose that we have a bijection from type i variables of the language of TST to type α variables of the language of TTT for each natural number i and ordinal $\alpha < \lambda$.

TTT is then the first order theory with types indexed by the ordinals below λ whose well formed atomic sentences '$x = y$' have $\mathtt{type}('x') = \mathtt{type}('y')$ and whose atomic sentences '$x \in y$' satisfy $\mathtt{type}('x') < \mathtt{type}('y')$, and whose axioms are the sentences ϕ^s for each axiom ϕ of TST and each strictly increasing sequence s in λ. TTTU has the same relation to TSTU (with the addition of constants $\emptyset^{\alpha,\beta}$ for each $\alpha < \beta < \lambda$ such that $(\forall \mathbf{x}_0^\alpha : \mathbf{x}_0^\alpha \notin \emptyset^{\alpha,\beta})$ is an axiom).

It is important to notice how weird a theory TTT is. This is not cumulative type theory. Each type β is being interpreted as a power set of *each* lower type α. Cantor's theorem in the metatheory makes it clear that most of these power set interpretations cannot be honest.

There is now a striking

Theorem (Holmes): TTT(U) is consistent iff NF(U) is consistent.

Proof: Suppose NF(U) is consistent. Let (M, E) be a model of NF(U) (a set M with a membership relation E). Implement type α as $M \times \{\alpha\}$ for each $\alpha < \lambda$. Define $E_{\alpha,\beta}$ for $\alpha < \beta$ as $\{((x, \alpha), (y, \beta)) : xEy\}$. This gives a model of TTT(U). Empty sets in TTTU present no essential additional difficulties.

Suppose TTT(U) is consistent, and so we can assume we are working with a fixed model of TTT(U). Let Σ be a finite set of sentences in the language of TST(U). Let n be the smallest type such that no type n variable occurs in any sentence in Σ. We define a partition of the n-element subsets of λ. Each $A \in [\lambda]^n$ is put in a compartment determined by the truth values of the sentences ϕ^s in our model of TTT(U), where $\phi \in \Sigma$ and $\mathtt{rng}(s \restriction \{0, \ldots, n-1\}) = A$. By Ramsey's theorem, there is a homogeneous set $H \subseteq \lambda$ for this partition, which includes the range of a strictly increasing sequence h. There is a complete extension of TST(U) which includes ϕ iff the theory of our model of TTT(U) includes ϕ^h. This extension satisfies $\phi \leftrightarrow \phi^+$ for each $\phi \in \Sigma$. But this implies by compactness that the full Ambiguity Scheme $\phi \leftrightarrow \phi^+$ is consistent with TST(U), and so that NF(U) is consistent by the 1962 result of Specker.

We note that we can give a treatment of the result of Specker (rather different from Specker's own) using TTT(U). Note that it is easy to see that if we have a model of TST(U) augmented with a Hilbert symbol (a primitive term construction $(\epsilon x : \phi)$ (same type as x) with axiom scheme $\phi[(\epsilon x : \phi)/x] \leftrightarrow (\exists x : \phi)$) which cannot appear in instances of comprehension (the quantifiers are not defined in terms of the Hilbert symbol, because they do need to appear in instances of comprehension) and Ambiguity (for all formulas, including those which mention the Hilbert symbol) then we can readily get a model of NF, by constructing a term model using the Hilbert symbol in the natural way, then identifying all terms with their type-raised versions. All statements in the resulting type-free theory can be decided by raising types far enough (the truth value of an atomic sentence $(\epsilon x : \phi) R (\epsilon y : \psi)$ in the model of NF is determined by raising the type of both sides (possibly by different amounts) until the formula is well-typed in TST and reading the truth value of the type raised version; R is either = or $\in$). Now observe that a model of TTT(U) can readily be equipped with a Hilbert symbol if this creates no obligation to add instances of comprehension containing the Hilbert symbol (use a well-ordering of the set implementing each type to interpret a Hilbert symbol $(\epsilon x : \phi)$ in that type as the first x such that ϕ), and the argument above for consistency of TST(U) plus Ambiguity with the Hilbert symbol goes through.

Theorem (essentially due to Jensen): NFU is consistent.

Proof: It is enough to exhibit a model of TTTU. Suppose $\lambda > \omega$. Represent type α as $V_{\omega+\alpha} \times \{\alpha\}$ for each $\alpha < \lambda$ ($V_{\omega+\alpha}$ being a rank of the usual cumulative hierarchy). Define $\in_{\alpha,\beta}$ for $\alpha < \beta < \lambda$ as

$$\{((x,\alpha),(y,\beta)) : x \in V_{\omega+\alpha} \wedge y \in V_{\omega+\alpha+1} \wedge x \in y\}.$$

This gives a model of TTTU in which the membership of type α in type β interprets each (y,β) with $y \in V_{\omega+\beta} \setminus V_{\omega+\alpha+1}$ as an urelement.

Our use of $V_{\omega+\alpha}$ enforces Infinity in the resulting models of NFU (note that we did not have to do this: if we set $\lambda = \omega$ and interpret type α using V_α we prove the consistency of NFU with the negation of Infinity). It should be clear that Choice holds in the models of NFU eventually obtained if it holds in the ambient set theory.

This shows in fact that mathematics in NFU is quite ordinary (with respect to stratified sentences), because mathematics in the models of TSTU embedded in the indicated model of TTTU is quite ordinary. The notorious ways in which NF evades the paradoxes of Russell, Cantor and Burali-Forti can be examined in actual models and we can see that they work and how they work (since they work in NFU in the same way they work in NF).

Of course Jensen did not phrase his argument in terms of tangled type theory. Our contribution here was to reverse engineer from Jensen's original argument for the consistency of NFU an argument for the consistency of NF itself, which requires additional input which we did not know how to supply (a proof of the consistency of TTT itself). An intuitive way to say what is happening here is that Jensen noticed that it is possible to skip types in a certain sense in TSTU in a way which is not obviously possible in TST itself; to suppose that TTT might be consistent is to suppose that such type skipping is also possible in TST.

2.4.1 How internal type representations unfold in TTT

We have seen above that TST can internally represent TST_n. An attempt to represent types of TTT internally to TTT has stranger results. The development of the model does not depend on reading this section.

In TST the strategy for representing type i in type $n \geq i$ is to use the $n - i$-iterated singleton of any type i object x to represent x; then membership of representations of type $i - 1$ objects in type i objects is represented by the relation on $n - i$-iterated singletons induced by the subset relation and with domain restricted to $n - (i+1)$-fold singletons. This is described more formally above.

In TTT the complication is that there are numerous ways to embed type α into type β for $\alpha < \beta$ along the lines just suggested. We define a generalized iterated singleton operation: where A is a finite subset of λ, ι_A is an operation defined on objects of type $\mathtt{min}(A)$. $\iota_{\{\alpha\}}(x) = x$. If A has $\alpha < \beta$ as its two smallest elements, $\iota_A(x)$ is $\iota_{A_1}(\iota_{\alpha,\beta}(x))$, where A_1 is defined as $A \setminus \{\mathtt{min}(A)\}$ (a notation we will continue to use) and $\iota_{\alpha,\beta}(x)$ is the unique type β object whose only type α element is x.

Now for any nonempty finite $A \subseteq \lambda$ with minimum α and maximum β. the range of ι_A is a set, and a representation of type α in type β. For simplicity we carry out further analysis in types $\beta, \beta+1, \beta+2 \ldots$ though it could be done in more general increasing sequences. Use the notation τ_A for the range of ι_A, for each set A with β as its maximum. Each such set has a cardinal $|\tau_A|$ in type $\beta + 2$. It is a straightforward argument in the version of TST with types taken from A and a small finite number of types $\beta+i$ that $2^{|\tau_A|} = |\tau_{A_1}|$ for each A with at least two elements. The relevant theorem in TST is that $2^{|\iota^{n+1}``X|} = \iota^n ``X$, relabelled with suitable types from λ. We use the notation $\exp(\kappa)$ for 2^κ to support iteration. Notice that for any τ_A we have $\exp^{|A|-1}(|\tau_A|) = |\tau_{\{\beta\}}|$, the cardinality of type β. Now if A and A' have the same minimum α and maximum β but are of different sizes, we see that $|\tau_A| \neq |\tau_{A'}|$, since one has its $|A|-1$-iterated exponential equal to $|\tau_{\{\beta\}}|$ and the other has its $|A'|-1$-iterated exponential equal to $|\tau_{\{\beta\}}|$. This is odd because there is an obvious external bijection between the sets τ_A and $\tau_{A'}$: we see that this external bijection cannot be realized as a set. τ_A and $\tau_{A'}$ are representations of the same type, but this is not obvious from inside TTT. We recall that we denote $A \setminus \{\mathtt{min}(A)\}$ by A_1; we further denote $(A_i)_1$ as A_{i+1}. Now suppose that A and B both have maximum β and $A \setminus A_i = B \setminus B_i$, where $i < |A| \leq |B|$. We observe that for any concrete sentence ϕ in the language of TST_i, the truth value of ϕ in natural models with base type of sizes $|\tau_A|$ and $|\tau_B|$ will be the same, because the truth values we read off are the truth values in the model of TTT of versions of ϕ in exactly the same types of the model (truth values of ϕ^s for any s having $A \setminus A_i = B \setminus B_i$ as the range of an initial segment). This much information telling us that τ_{A_j} and τ_{B_j} for $j < i$ are representations of the same type is visible to us internally, though the external isomorphism is not. We can conclude that the full first-order theories of natural models of TST_i with base types $|\tau_A|$ and $|\tau_B|$ are the same as seen inside the model of TTT, if we assume that the natural numbers

of our model of TTT are standard.

2.4.2 Tangled webs of cardinals: a suggestion of another approach not followed here

Nothing in the construction of a model of tangled type theory and verification that it is a model which appears below depends on anything in this section.

It is straightforward to transform a model of TST into a model of bounded Zermelo set theory (Mac Lane set theory) with atoms or without foundation (this depends on how type 0 is handled). Specify an interpretation of type 0 either as a set of atoms or a set of self-singletons. Then interpret type $i + 1$ as inhabited by sets of type i objects in the obvious way, identifying type $i + 1$ objects with objects of lower type which happen to have been assigned the same extension.

In a model of TTT, do this along some increasing sequence of types of order type ω whose range includes an infinite ordinal α. In the resulting model of bounded Zermelo set theory, let τ_A represent the cardinality of the range of ι_A as in the previous discussion (for nonempty subsets of type A all with maximum the same infinite α). Suppose further for the sake of argument that our model of TTT is λ-complete, in the sense that any subset of a type of cardinality that of λ or less is implemented as a set in each higher type. It will follow that $A \mapsto \tau_A$ is actually a function. [It is an incidentally interesting fact that the models we construct (with no dependence on this section) actually have this completeness property].

We describe the situation which holds for these cardinals.

We work in Mac Lane set theory. Choice is not assumed, and we use the Scott definition of cardinals.

Definition: If A is a nonempty finite set of ordinals which is sufficiently large, we define A_1 as $A \setminus \mathtt{min}(A)$ and A_0 as A, A_{i+1} as $(A_i)_1$.

Definition: A tangled web of cardinals of order α (an infinite ordinal) is a function τ from the set of nonempty sets of ordinals with α as maximum to cardinals such that

1. If $|A| > 1$, $\tau(A_1) = 2^{\tau(A)}$.

2. If $|A| \geq n$, the first order theory of a natural model of TST_n with base type $\tau(A)$ is completely determined by $A \setminus A_n$, the n smallest elements of A.

The bookkeeping in different versions of this definition in different attempts at a tangled web version of the proof of the consistency of NF have been different (an obvious point about the version given here is that the top ordinal α could be omitted). Another remark is that it is clear that asserting the existence of a tangled web is stronger than simple TTT: it requires $\lambda > \omega$, and the λ-completeness of course is a strong assumption in the background. All variants that I have used support versions of the following

Theorem: If there is a model of Mac Lane set theory in which there is a tangled web of cardinals τ, then NF is consistent.

Proof: Let Σ be a finite set of sentences of the language of TST. Let n be larger than any type mentioned in any formula in Σ. Partition $[\alpha]^n$ into compartments in such a way that the compartment that a set A is put into depends on the truth values of the sentences in Σ in natural models of TST_n with base type of size $\tau(B)$ where $B \setminus B_n = A$. This partition of $[\alpha]^n$ into no more than $2^{|\Sigma|}$ compartments has a homogeneous set H of size $n + 1$. The natural models of TST_n with base types of size $\tau(H)$ and base types of size $\tau(H_1)$ have the same truth values for sentences in Σ, so the model of TST with base type $\tau(H)$ satisfies the restriction of the Ambiguity Scheme to Σ, so the full Ambiguity Scheme is consistent by compactness, so TST + Ambiguity is consistent so NF is consistent.

Our initial approach to proving our theorem was to attempt a Frankel-Mostowski construction of a model of Mac Lane set theory with a tangled web of cardinals. We do know how to do this, but we believe from recent experience that constructing a model of tangled type theory directly is easier, though tangled type theory is a nastier theory to describe.

We think there is merit in giving a brief description of a situation in a more familiar set theory equivalent to (a strengthening of) the very strange situation in a model of tangled type theory. This section is also useful here because it supports the discussion in the conclusion of one of the unsolved problems which is settled by this paper.

2.5 An axiomatization of TST with finitely many templates

We discuss a finite axiomization of NF derived from that of Hailperin (it is taken from an implementation of Hailperin's axiom set of [4] in Metamath ([2]), and there are minor changes from the original formulation), making the important observation that it actually provides us with an axiomatization of TST with finitely many axiom templates (in the sense that each axiom is a type shifted version of one of a finite set of axioms). Notation introduced in this section is not used in the rest of the paper, and nothing in subsequent sections except a brief remark in the last paragraphs of section 4 depends on anything here. This finite axiomatization is however used in the Lean formalization.

The finite axiomatization of NF takes this form (the definitions inserted are ours, and we have modified the order of the axioms to make the definitions work sensibly). We also present this as an axiomatization of TST with finitely many templates, with the proviso that each typed form of each axiom is asserted:

extensionality axiom: $(\forall x : (\forall y : (\forall z : z \in x \leftrightarrow z \in y) \to x = y))$

anti-intersection axiom: $(\forall xy : (\exists z : (\forall w : w \in z \leftrightarrow \neg(w \in x \land w \in y))))$

singleton axiom: $(\forall x : (\exists y : (\forall z : z \in y \leftrightarrow z = x)))$

definition: $\{x\}$ denotes for each x the set whose only element is x, whose existence is provided by the previous axiom. We define $\iota(x)$ as $\{x\}$ and define $\iota^1(x)$ as $\iota(x)$ and $\iota^{n+1}(x)$ as $\{\iota^n(x)\}$, for each concrete natural number n.

cardinal one axiom: $(\exists x : (\forall y : y \in x \leftrightarrow (\exists z : (\forall w : w \in y \leftrightarrow w = z))))$

definition: We define 1 as the set of all singletons, provided by the previous axiom.

definition: $x|y$ denotes the set z whose existence is provided by the anti-intersection axiom: $z \in x|y \leftrightarrow \neg(z \in x \land z \in y)$. We define x^c as $x|x$. We define $x \cap y$ as $(x|y)^c$. We define $x \cup y$ as $x^c|y^c$. We define V as $1|1^c$ (noting that it is straightforward to prove $x|x^c = V$ for any x, since this is the universal set). We define $\{x, y\}$ as $\{x\} \cup \{y\}$. We define (x, y) as $\{\{x\}, \{x, y\}\}$. We define (x, y, z) as $(\{\{x\}\}, (y, z))$. More generally, we define $(x_1, \ldots, x_n)$ as $(\iota^{2n-4}(x_1), (x_2, \ldots, x_n))$ for $n > 2$ a concrete natural number. [The treatment of n-tuples is what makes this axiomatization singularly awful].

cross product axiom: $(\forall x : (\exists y : (\forall z : z \in y \leftrightarrow (\exists wt : z = (w, t) \land t \in x))))$

definition: We define $V \times x$ as the set introduced by the previous axiom: $z \in V \times x \leftrightarrow (\exists wt : z = (w, t) \land t \in x)$. Note that $V \times V$ is the set of all ordered pairs.

converse axiom: $(\forall x : (\exists y : (\forall zw : (z, w) \in y \leftrightarrow (w, z) \in x)))$

definition: For any set R, we define R^{-1} as the intersection of $V \times V$ with a set introduced by the previous axiom:

$$(\forall zw : (z,w) \in R^{-1} \leftrightarrow (w,z) \in R)) \wedge (\forall u : u \in R^{-1} \leftrightarrow (\exists zw : (z,w) = u)).$$

definition: We define $x \times V$ as $(V \times x)^{-1}$ and $x \times y$ as $(x \times V) \cap (V \times y)$.

singleton image axiom: $(\forall x : (\exists y : (\forall zw : (\{z\}, \{w\}) \in y \leftrightarrow (z,w) \in x)))$.

definition: We define R^{ι} for any set R as the intersection of a set provided by the previous axiom with 1×1. R^{ι} is the set whose members are exactly the ordered pairs $(\{z\}, \{w\})$ such that $(z,w) \in R$. Let R^{ι^1} be defined as R^{ι} and $R^{\iota^{n+1}}$ be defined as $(R^{\iota^n})^{\iota}$.

We define $\iota^2 \text{``} x$ as $(x \times x) \cap V$. This is the image of x under the double singleton operation.

Note that $\iota^2 \text{``} V$ is the equality relation. Define $\iota^{2(n+1)} \text{``} x$ as $\iota^2 \text{``}(\iota^{2n} \text{``} x)$.

We define $x_1 \times x_2 \ldots \times x_n$ as $\iota^{2(n-2)} \text{``} x \times (x_2 \times \ldots \times x_n)$.

insertion two axiom: $(\forall x : (\exists y : (\forall zwt : (z,w,t) \in y \leftrightarrow (z,t) \in x)))$

We define $I_2(R)$ as the intersection of a set provided by the previous axiom with $V \times V \times V$.

insertion three axiom: $(\forall x : (\exists y : (\forall zwt : (z,w,t) \in y \leftrightarrow (z,w) \in x)))$

We define $I_3(R)$ as the intersection of a set provided by the previous axiom with $V \times V \times V$.

definition: It seems natural to define $I_1(R)$ as $\iota^2 \text{``}(V) \times R$, but this requires no new axiom.

definition: Define $I_{1,n}(R)$ as $\iota^{2(n-2)} \text{``} V \times R$: this is correct for prepending all possible initial projections to an n-tuple. Then define $I_{1,n}^1(R)$ as $I_{1,n}(R)$ and define $I_{1,n}^{m+1}(R)$ as $I_{1,n+m+1}(I_{1,n}^m(R))$: this takes into account the fact that the tuples get longer.

Define $I_{2,n}(R)$ as $I_2(R) \cap (V \times \iota^{2(n-3)}(V) \times V)$, and define $I_{2,n}^{m+1}(R)$ as $I_{1,n+m+1}(I_{2,n}^m(R))$: this takes into account the fact that the tuples get longer.

type lowering axiom: $(\forall x : (\exists y : (\forall z : z \in y \leftrightarrow (\forall w : (w, \{z\}) \in x))))$

definition: We define $\mathtt{TL}(x)$ by $(\forall z : z \in \mathtt{TL}(x) \leftrightarrow (\forall w : (w, \{z\}) \in x))$. This is a very strange operation!

We define $\iota^{-1} \text{``} x$ as $\mathtt{TL}(V \times x)$. This is the set of all elements of singletons belonging to $x$. We can then define $\iota \text{``} x$, the elementwise image of x under the singleton operation, as $\iota^{-1} \text{``}(\iota^2 \text{``}(x))$.

Further, we define $\iota^{-(n+1)} \text{``}(x)$ as $\iota^{-1} \text{``}(\iota^{-n} \text{``}(x))$ for each concrete natural number $n$, and $\iota^{n+1} \text{``}(x)$ as $\iota \text{``}(\iota^n \text{``}(x)$

We develop an important operation step by step.

$\mathtt{TL}(R) = \{z : (\forall w : (w, \{z\}) \in R)\}$

Dually, $(\mathtt{TL}(R^c))^c = \{z : (\exists w : (w, \{z\}) \in R)\}$

Now $(\mathtt{TL}((R^\iota)^c))^c = \{z : (\exists w : (w, \{z\}) \in R^\iota)\}$,

which is the same as $(\mathtt{TL}((R^\iota)^c))^c = \{z : (\exists w : (\{w\}, \{z\}) \in R^\iota)\}$ because all elements of the domain of R^ι are singletons,

which is the same as $(\mathtt{TL}((R^\iota)^c))^c = \{z : (\exists w : (w, z) \in R)\}$

so we define $\mathtt{rng}(R)$ as $(\mathtt{TL}((R^\iota)^c))^c$, and define $\mathtt{dom}(R)$ as $\mathtt{rng}(R^{-1})$.

subset axiom: $(\exists x : (\forall yz : (y, z) \in x \leftrightarrow (\forall w : w \in y \rightarrow w \in z)))$

We define $[\subseteq]$ as the intersection of a set provided by the previous axiom with $V \times V$: $[\subseteq]$ is the set of all ordered pairs (x, y) such that $x \subseteq y$.

We define $[\in]$ as $[\subseteq] \cap 1 \times V$.

This is not our favorite finite axiomatization of NF (or our favorite finite template axiomatization of TST) but it is the one in the formalization and also basically the oldest one, so we present a verification of it in outline at least.

What we need to do is verify that $\{x : \phi\}$ exists for each formula ϕ of the language of TST, to ensure that comprehension holds. We do this by induction on the structure of formulas.

$\{x : \neg\phi\}$ is $\{x : \phi\}^c$.

$\{x : \phi \wedge \psi\}$ is $\{x : \phi\} \cap \{x : \psi\}$.

Now we have the much more complex task of analyzing

$$\{t_n : (\forall x_i : \phi(t_1, \ldots, t_n)\}.$$

Choose a type τ' higher than the type τ_i of each t_i. Do this for bound variables as well, and further, in each formula $(\forall t_i : \psi)$ or $(\exists t_i : \psi)$ we require that all occurrences of t_i be free in ψ and the index of t_i be less than the index of any other variable appearing free in ψ. It should be clear that we can do this without loss of generality.

Where the type of t_i is τ_i, we define x_i as $\iota^{\tau'-\tau_i}(t_i)$: we construct

$$\{t_n : (\forall x_i : \phi(t_1, \ldots, t_n)\}$$

by defining manipulations which allow us to build sets $\{(x_1, \ldots, x_n) : \phi^*(x_1, \ldots, x_n)\}$ in which all the variables are of the same type. We write ϕ^* to suggest that the formula ϕ must be transformed to effect our change of variables: $t_i = t_j$ is equivalent to $x_i = x_j$, and $t_i \in t_j$ is equivalent to $(x_i.x_j) \in [\epsilon]^{\iota^{\tau'-\tau_j}}$ (the reader will see that we use this representation below, though embedded in larger tuples). A quantifier over t_i is equivalent to a quantifier over x_i restricted to $\iota^{\tau'-\tau_i}{}^{\shortmid\shortmid}V$.

For any such representation, we have a type signature

$$\iota^{\tau'-\tau_1}{}^{\shortmid\shortmid}V \times \ldots \times \iota^{\tau'-\tau_n}{}^{\shortmid\shortmid}V.$$

We abbreviate this as τ^*.

The set $\{(x_1,\dots,x_n) : x_1 = x_n\}$ is obtained as $I_{2,2}^{n-2}(\iota\,"V) \cap \tau^*$.

The set $\{x_1,\dots,x_n) : x_1 = x_i\}$ $(i < n)$ is obtained as $I_{2,3}^{i-3}(I_3(\iota\,"V)) \cap \tau^*$.

We then can represent the set $\{(x_1,\dots,x_n) : x_i = x_j\}$ (wlog $i < j$) as $I_{1,n-i+1}^{i-1}\{(x_1,\dots,x_{n-i+1}) : x_1 = x_{j-i+1}\}) \cap \tau^*$.

The set $\{(\{x_1\}, x_2) : x_1 \in x_2\}$ has already been defined above as $[\in]$.

The set $\{(x_1,\dots,x_n) : (\exists uv : x_i = \iota^{k+1}(u) \wedge x_j = \iota^k(v) \wedge u \in v)\}$ is

$$\mathtt{rng}^2(I_3((R^{\iota^{k+2n}}[\in]) \cap \{(x_1,\dots,x_{n+2} : x_1 = x_{i+2}\} \cap \{(x_1,\dots,x_{n+2}) : x_2 = x_{j+2}\}).$$

This handles all representations of membership statements between x_i's in the framework we are using.

The set $\{(x_1,\dots,x_n) : (\exists x_1 : (x_1,\dots,x_n) \in R)\}$ is representable as $I_{1,n-1}(\mathtt{dom}(R^{-1}))$.

This only allows us to quantify over the lowest numbered variable. This is actually sufficient. We have renamed bound variables so that all bound variables with different binders are distinct and in any subformula $(\exists y : \psi)$ y will have lower index than any variable free in ψ. Then the quantified variable might not be x_1, but if it is x_k, we can use the representation $I_{1,n-k}^k(\mathtt{dom}(\mathtt{rng}^{k-1}(R))^{-1})$: strip off the first $k - 1$ variables, which are dummies already quantified over, quantify over x_k and put k dummies back, as it were.

This gives sufficient machinery to handle representations of all sentences in the language of TST (or stratified sentences in the language of NF) in the format we are using. However, the final representation of $\{t_n : \phi(t_1,\dots,t_n)\}$ obtained in this way will be of the form $R = \{(\iota^{\tau'-\tau_1}(t_1),\dots,\iota^{\tau'-\tau_n}(t_n)) : \phi(t_1\dots,t_n)\}$ where some of the t_i's are bound variables (all values of which will appear), and some are parameters. Let P be the set of i such that t_i is a parameter.

The final computation of $\{t_n : \phi(t_1,\dots,t_n)\}$ will be as

$$\iota^{-(\tau'-\tau_n)}\,"(\mathtt{rng}^{n-1}(R \cap \bigcap_{i \in P}\{(x_1,\dots,x_n) : x_i = \iota^{\tau_i-\tau}(t_i)\}))$$

where a final bit of computation must be exhibited: $\{(x_1,\dots,x_n) : x_i = t\}$ is realized as $I_{1,2}^{i-1}(\{\iota^{2(n-i-1)}(t_i)\} \times V)) \cap \tau^*$.

We state for the record that we think this is a bad finite axiomatization of NF, or finite template axiomatization of TST: we think the treatment of n-tuples is terribly difficult to work with. But it does work, and the second author chose to verify it, so it merits discussion here. It should be noted that any finite template axiomatization of TST could be used; there is no advantage to this one for the formalization, and there is an oddity, because the one impredicative axiom (the axiom of type lowering) does a lot of extra work, since it is also essential for constructing domains and ranges of functions. In the formalization, the proof of type lowering was in effect divided into the proof of the existence of $\iota^3\,"\mathtt{TL}(x)$, which is predicative, and the proof of the existence of set unions of sets of singletons.

3   The model description

In this section we give a complete description of what we claim is a model of tangled type theory. Our metatheory is some fragment of ZFC.

3.1   The abstract supertype framework

Abstract considerations about types; the system of supertypes defined:

Types are indexed by a well-ordering $\leq_\tau$ (from which we define a strict well-ordering $<_\tau$ in the obvious way). We refer to elements of the domain of $\leq_\tau$ as "type indices".

We first define a system of "supertypes" (using the same type labels).

For each element t of $\mathtt{dom}(\leq_\tau)$ we will define a set τ_t^*, called supertype t.

If m is the minimal element of the domain of $\leq_\tau$, we choose a set τ_m^* as supertype m.

$\leq_\tau$ and τ_m^* are the only parameters of the system of supertypes (which is not a model of TTT, but a sort of maximal structure for the language of TTT).

We describe the construction of τ_t^*, assuming that $t \in \mathtt{dom}(\leq_\tau)$ and $t \neq m$ and for all $u <_\tau t$, we have defined τ_u^*.

We define τ_t^* as

$$\{X \cup \{\{\tau_u^* : u <_\tau t\}\} : X \subseteq \bigcup_{u <_\tau t} \tau_u^*\}.$$

An element of supertype $t >_\tau m$ is a subset of the union of all lower types, with $t^+ = \{\tau_u^* : u <_\tau t\}$ added as an element.

Foundation in the metatheory ensures a clean construction here. An element x of supertype $t >_\tau m$ is always nonempty with t^+ as an element. The set t^+ has supertype u as an element for each $u <_\tau t$, so t^+ and so x cannot belong to any supertype u with $u <_\tau t$, by foundation: each other element of x belongs to such a supertype. We have shown that all the types are disjoint. The labelling element t^+ cannot belong to supertype t by foundation, because an element of supertype t must be nonempty and have t^+ as an element. Further, t^+ cannot belong to any supertype v with $t <_\tau v$, because any element of v contains v^+ as an element which contains supertype t as an element and any element of supertype t contains t^+ as an element, so $t^+ \in v$ would violate foundation in the metatheory.

The membership relations of this structure are transparent: $x \in_{t,u} y$ ($t <_\tau u$) is defined as $x \in \tau_t^* \wedge y \in \tau_u^* \wedge x \in y$. Considerations above show that there are no unintended memberships caused by the labelling elements t^+, because the labelling elements cannot themselves belong to any supertype. Note the presence of $\emptyset_t = \{t^+\}$ in supertype t, which has no elements of any type $u <_\tau t$ (and is distinct from $\emptyset_v$ for $v \neq t$).

The system of supertypes is certainly not a model of TTT, because it does not satisfy extensionality. It is easy to construct many sets in a higher type with the same extension over a given lower type, by modifying the other extensions of the object of higher type.

The system of supertypes does satisfy the comprehension scheme of TTT. One can use Jensen's method to construct a model of stratified comprehension with no extensionality axiom from the system of supertypes, and stratified comprehension with no extensionality axiom interprets NFU in a manner described by Marcel Crabbé in [1].

the generality of the system of supertypes: We show that any model of TTT (assuming there are any) will be in effect isomorphic to a substructure of a system of supertypes.

Let M be a model of TTT (more generally, any structure for the language of TTT in which each object is determined given all of its extensions). Let $\leq_M$ be the well-ordering on the types of M and let m be the minimal type of M. We will assume as above that $\leq_\tau$ is a well-ordering of type labels t with corresponding actual types τ_t of M: of course, we could use the actual types of M as type indices, but we preserve generality this way. We also assume that the sets implementing the types of M are disjoint (it is straightforward to transform a model in which the sets implementing the types are not disjoint to one in which they are, without disturbing its theory, by replacing each $x \in \tau_t$ with (x, t)).

We consider the supertype structure generated by $\leq_\tau := \leq_M$ and $\tau_m^* := \tau_m$. We indicate how to define an embedding from M into this supertype structure.

Define $I(x) = x$ for $x \in \tau_m = \tau_m^*$.

If we have defined I on each type $u <_\tau t$, we define $I(x)$, for $x \in \tau_t$, as $\bigcup_{u <_\tau t} \{I(y) : y \in_{u,t}^M x\} \cup \{\{\tau_u^* : u <_\tau t\}\}$.

It should be clear that as long as M satisfies the condition that an element of any type other than the base type is uniquely determined given all of its extensions in lower types, I is an isomorphism from M to a substructure of the stated system of supertypes. A model of TTT, in which any one extension of an element of any higher type in a lower type exactly determines the object of higher type, certainly satisfies this condition. So the problem of constructing a model of TTT is equivalent to the problem of constructing a model of TTT which is a substructure of a supertype system.

Some advantages of this framework are that the membership relations in TTT are interpreted as subrelations of the membership relation of the metatheory, while the types are sensibly disjoint.

Note on general systems of types and extensions therein: Notice that if $\alpha <_\tau \beta$ are supertypes, and $x \in \tau_\alpha$, $x \cap \tau_\beta^*$ is the extension of x over supertype β. This will be inherited by a scheme of types τ_γ with each $\tau_\gamma \subseteq \tau_\gamma^*$

if an additional condition holds: for $\alpha > \beta$, we will have for $x \in \tau_\alpha$ that $x \cap \tau_\beta^*$ is the extension of x over supertype β as already noted: for it to be the extension over type β we need the general condition $x \cap \tau_\gamma^* \subseteq \tau_\gamma$ for all $\delta > \gamma$ and $x \in \tau_\delta$.

3.2 Preliminaries of the construction

preliminaries of our construction; cardinal parameters and type -1:
Now we introduce the notions of our particular construction in this framework.

Let λ be a limit ordinal. Let $\leq_\tau$ be the order on $\lambda \cup \{-1\}$ which has -1 as minimal and agrees otherwise with the usual order on λ.

Let $\kappa > \lambda$ be a regular cardinal (that is, a regular initial ordinal). Note that since $\kappa > \lambda$, κ is uncountable. Sets of cardinality $< \kappa$ we call "small". Sets which are not small we may call "large".

Let μ be a strong limit cardinal greater than κ with cofinality at least κ.

Let $\tau_{-1}^* = \tau_{-1}$ be

$$\{(\nu, \beta, \gamma, \alpha) : \nu < \mu \wedge \beta \in \lambda \cup \{-1\} \wedge \gamma \in \lambda \setminus \{\beta\} \wedge \alpha < \kappa\}.$$

Note that this completes the definition of the supertype structure we are working in: we now have a definite reference for τ_α^* for $\alpha \in \lambda$.

type shorthand, extended type indices: Notice that if α, β are type indices, $\alpha \in \beta$ is a convenient short way to say $-1 <_\tau \alpha <_\tau \beta$. A nonempty finite subset of $\lambda \cup \{-1\}$ may be termed an *extended type index*. If A is an extended type index with at least two elements, A_1 is defined as $A \setminus \{\mathtt{min}(A)\}$.

atoms, litters and near-litters: We may refer to elements of τ_{-1} (or closely related objects) as "atoms" from time to time, though they are certainly not atomic in terms of the metatheory.

A *litter* is a subset of τ_{-1} of the form $L_{\nu,\beta,\gamma} = \{(\nu, \beta, \gamma, \alpha) : \alpha < \kappa\}$. The litters make up a partition of type -1 (which is of size μ) into size κ sets.

On each litter $L = L_{\nu,\beta,\gamma}$ define a well-ordering $\leq_L$: $(\nu, \beta, \gamma, \alpha) \leq_L (\nu, \beta, \gamma, \alpha')$ iff $\alpha \leq \alpha'$. The strict well-ordering $<_L$ is defined in the obvious way. This well-ordering is used in only one place in the paper, and its use could easily be avoided, but we find its concreteness appealing.

A *near-litter* is a subset of τ_{-1} with small symmetric difference from a litter. We define $M \sim N$ as $|M \Delta N| < \kappa$, for M, N near-litters: in English, we say "M is near N" iff $M \sim N$. Note that nearness is an equivalence relation on near-litters.

We define N°, for N a near-litter, as the litter L such that $L \sim N$.

cardinality of the collection of near-litters :

We verify that there are μ near-litters.

A near-litter is determined as the symmetric difference of a litter L (there are μ litters) and a small subset of τ_{-1}.

So it is sufficient to show that a set of size μ has no more than μ small subsets. We note that μ is a strong limit cardinal of cofinality $\geq \kappa$. We define a sequence of cardinals and a sequence of functions f indexed by ordinals. $\mu_0 = 0$. f_0 is the empty function. $\mu_{\alpha+1} = 2^{\mu_\alpha}$. $f_{\alpha+1}$ is a bijection to 2^{μ_α} from the smallest segment in μ of the correct cardinality whose lower limit is the first ordinal not in the range of f_α. For a limit ordinal λ, μ_λ is the supremum of the set of μ_β for $\beta < \lambda$ and f_λ is a bijection to 2^{μ_λ} from the smallest segment in μ of the correct cardinality whose lower limit is the first ordinal not in the range of any f_β for $\beta < \lambda$. Because μ is a strong limit cardinal of cofinality $\geq \kappa$, the first undefined f_α must be at a limit and satisfy $\alpha \geq \kappa$ and $\mu_\alpha = \mu$. The union of the maps f_α is a function with domain μ whose range includes every small subset of μ, because any small subset of μ must be included in some μ_α, $\alpha < \mu$.

3.3 Machinery for enforcing extensionality in the model

We describe the mechanism which enforces extensionality in the substructure of this supertype structure that we will build.

The levels of the structure we will define are denoted by τ_α for $\alpha \in \lambda \cup \{-1\}$. As we have already noted, $\tau_{-1} = \tau_{-1}^*$ as defined above.

In defining $\tau_\alpha \subseteq \tau_\alpha^*$ for each α, we assume that we have already defined τ_β for each $\beta < \alpha$, and that the system of types $\{\tau_\beta : \beta <_\tau \alpha\}$ already defined satisfies various hypotheses which we will discuss as we go. Elements x of τ_α^* which we consider for membership in τ_α will have $x \cap \tau_\beta^* \subseteq \tau_\beta$ for $\beta < \alpha$. We assume that for $\gamma < \beta < \alpha$, if $x \in \tau_\beta$, $x \cap \tau_\gamma^* \subseteq \tau_\gamma$.

We suppose that each τ_β already constructed is of cardinality μ. Note that we already know that τ_{-1} is of cardinality μ.

We further intimate that for each $x \in \tau_\gamma$, $-1 < \gamma < \alpha$, we have defined objects S for which we say that S is a support of x. The definition of supports will be given later. For the moment, we define τ_γ^+ as the set of all (x, S) for which $x \in \tau_\gamma$ and S is a support of x. It is a hypothesis of the recursion that τ_γ is of cardinality μ, from which it follows that τ_γ^+ is of cardinality μ, since there are μ supports (as we will see when supports are defined). We provide that τ_{-1}^+ is the set of all $(x, \emptyset)$ for $x \in \tau_{-1}$.

We also provide a well-ordering $\leq_\gamma^+$ of order type μ of τ_γ^+ ($-1 \leq \gamma < \alpha$), by postulating an injection ι_γ^+ from τ_γ^+ into μ (it does not need to be onto) and defining $x \leq_\gamma^+ y$ as $\iota_\gamma^+(x) \leq \iota_\gamma^+(y)$. There are some hypotheses of the recursion about maps ι_γ^+ which are stated below. For any model element x and support S of x, we define $\iota_*^+(x, S)$ as $\iota_\gamma^+(x, S)$, where $x \in \tau_\gamma$.

We provide that for every near litter N and every $\beta < \alpha$, there is a unique element N_β of τ_β such that $N_\beta \cap \tau_{-1} = N$ (we will quite shortly give a precise description of all extensions of this object).

Definition: If X is a subset of type γ and $\gamma < \beta$, we define X_β as the unique element Y of τ_β such that $Y \cap \tau_\gamma = X$ (if this exists). Of course, this notation is only usable to the extent that we suppose that extensionality holds. Notice that the notation N_β is a case of this.

If $x \in \tau_{-1}$, we refer to any $\{x\}_\beta$ as a *typed atom* (of type β) and if N is a near-litter we refer to N_β as a *typed near-litter* (of type β).

We further stipulate that extensionality holds for each $\beta \in \alpha$ (for each $\gamma \in \beta$, any $x \in \tau_\beta$ is uniquely determined by $x \cap \tau_\gamma$; x is uniquely determined by $x \cap \tau_{-1}$ only on the additional assumption that $x \cap \tau_{-1}$ is nonempty)

position function for atoms and near-litters: We posit a bijection ι_* from the set of all atoms and near-litters to μ with the following properties:

1. $\iota_*(L) < \iota_*(a)$ if $a \in L$ and L is a litter.
2. $\iota_*(N^\circ) < \iota_*(N)$ if N is a near-litter which is not a litter
3. $\iota_*(a) < \iota_*(N)$ if $a \in N \Delta N^\circ$

The function may be constructed directly: first choose an ordering of type μ on litters, then put atoms directly after the litter they are inside (and before all later litters), then put near-litters after the corresponding litter and all atoms in the symmetric difference, then map each atom or near-litter to its position in the resulting order. We don't run out of room before finishing the construction because μ has cofinality at least κ.

An additional property involving ι_* which is enforced by inductive hypotheses explained later about the maps ι_β^+ is that $\iota_*(N) \leq \iota_*^+(N_\delta, S)$ will hold for any near litter N and support S of N_δ. It must be noted here because it is shortly used.

the definition of the f maps (crucially important) [1]:

We define for each type index β and each ordinal γ less than α and distinct from β a function $f_{\beta,\gamma}$ (whose definition does not actually depend on α: it will be the same at every stage). $f_{\beta,\gamma}$ is an injection from τ_β^+ into the set of litters, with range included in $\{L_{\nu,\beta,\gamma} : \nu < \mu\}$ to ensure that distinct f maps have disjoint ranges.

When we define $f_{\beta,\gamma}(x)$, we presume that we have already defined it for $y <_\beta^+ x$. We define $f_{\beta,\gamma}(x)$ as L, where $\iota_*(L)$ is minimal such that

---

[1] The use of model elements with support rather than simply model elements as domain elements of the f maps is a substantial contribution of the second author to the mathematics of the paper, above simply verifying the work of the first author. The proof could be carried out without this, but it is much easier to present with this refinement. There are other ways in which the second author has contributed to the mathematics, but this one is especially worthy of note.

1. $L \in \{L_{\nu,\beta,\gamma} : \nu < \mu\}$,
2. $\iota_*^+(x) < \iota_*(L)$ [and so for any $N \sim L$, $\iota_*^+(x) < \iota_*(N)$, and for any $z \in L$, $\iota_*^+(x) < \iota_*(z)$],
3. and for any $y <_\beta^+ x$, $f_{\beta,\gamma}(y) \neq L$.

Note that the ranges of distinct f maps are disjoint sets.

Definition (pre-extensional, distinguished extension): We define the notion of *pre-extensional* element of τ_β^* $(-1 < \beta \leq \alpha)$. An element x of τ_β^* is pre-extensional iff there is a $\gamma < \beta$ such that (1) $x \cap \tau_\gamma^* \subseteq \tau_\gamma$, and (2) $\gamma = -1$ if $x \cap \tau_{-1}$ is nonempty or if any $x \cap \tau_\delta$ $(\delta \in \beta)$ is empty, and (3) for each $\delta \in \beta \setminus \{\gamma\}$,

$$x \cap \tau_\delta = \{N_\delta : (\exists a \in x \cap \tau_\gamma : \exists S : N \sim f_{\gamma,\delta}(a, S))\}.$$

We say for any $x \in \tau_\beta^*$ and γ with this property that $x \cap \tau_\gamma$ is a *distinguished extension* of x.

We presume that all elements of τ_β, $\beta \in \alpha$, are pre-extensional.

Note that we now know how to compute all other extensions of typed near-litters N_β, because the -1-extension of N_β is distinguished and this indicates how to compute all the other extensions.

Verification that distinguished extensions are unique: We verify that the order conditions in the definition of $f_{\beta,\gamma}$ ensure that for any $x \in \tau_\beta$ there is only one set $x \cap \tau_\gamma$ which is a distinguished extension of x. If any $x \cap \tau_\gamma$ $(\gamma \in \beta)$ is empty or if $x \cap \tau_{-1}$ is not empty, $x \cap \tau_{-1}$ is the unique distinguished extension (if it is empty of course it coincides with all the other extensions).

So, what remains is the case of x with $x \cap \tau_{-1}$ empty and each $x \cap \tau_\gamma$ nonempty for $\gamma < \beta$.

If $x \cap \tau_\gamma$ is an extension of x and $a \in x \cap \tau_\gamma$ is not of the form N_δ for a near-litter N, then $x \cap \tau_\gamma$ must be the unique distinguished extension: the reason for this is that for any distinguished extension, all elements of other extensions must be of this form.

So we are down to the case where $x \cap \tau_\delta$ is nonempty for $\delta \neq -1$, $x \cap \tau_{-1}$ is empty, and each element of any $x \cap \tau_\delta$ is of the form N_δ where N is a near-litter. Let $x \cap \tau_\gamma$ be a distinguished extension. For any $M_\delta \in x$ with $\delta \neq \gamma$, we have $M \sim f_{\gamma,\delta}(P_\gamma, Q)$ for some near-litter P and support Q of P_γ. Then we have $\iota_*(M) > \iota_*^+(P_\gamma, Q) \geq \iota_*(P)$. Let $N_\gamma \in x$ be chosen so that $\iota_*(N)$ is minimal. It follows that for any $\delta \neq \gamma$ and $M_\delta \in x$ we have $\iota_*(M) > \iota_*(N)$, from which it is evident that we cannot have two distinct distinguished extensions: if δ were the index of another distinguished extension, and $\iota_*(M)$ were chosen minimal so that $M_\delta \in x$, it would follow that $\iota(M) < \iota(N)$ as above but also that $\iota(N) < \iota(M)$, which is absurd.

Definition (A maps and A^{-1}): For any $\delta \in \alpha$ and nonempty subset a of type $\gamma \neq \delta$, we define $A_\delta(a)$ as

$$\{N_\delta : (\exists x \in a : \exists S : N \sim f_{\gamma,\delta}(x, S))\}.$$

For any nonempty subset of type δ there is at most one subset y of any type such that $A_\delta(y) = x$. There cannot be more than one such y in any given type because the f maps are injective. There cannot be more than one such y in different types because the ranges of f maps with distinct index pairs are disjoint. We use the notation $A^{-1}(y)$ for this set if it exists, defining a very partial function A^{-1} on nonempty subsets of types.

Note that the distinguished extension of any type element x is the image under A^{-1} of the other extensions.

No subset of a type has infinitely many iterated images under A^{-1}: Let a be a subset of type γ for which $A^{-1}(a)$ exists.

Since $A^{-1}(a)$ exists, every element of a is of the form N_γ where N is a near-litter. Choose N such that $\iota_*(N)$ is minimal. Note that N is in fact a litter.

Let $b = A^{-1}(a)$: we have $a = A_\gamma(b)$, where b is a subset of some τ_δ with $\delta \neq \gamma$. In particular, $N = f_{\delta,\gamma}(u, U)$ for some $u \in \tau_\delta$ and U a support of u. If $u = M_\delta$ we can further state that $\iota_*(N) > \iota_*^+(M_\delta, U) \geq \iota_*(M)$: if b is itself an image under A^{-1}, the minimum value of ordinals $\iota_*(M)$ for $M_\delta \in b$ will be less than the minimum value of ordinals $\iota_*(N)$ for $N_\gamma \in a$, which establishes that there is an ordinal parameter determined by a nonempty subset of a type which decreases strictly when A^{-1} is applied (if it is applicable), and so no nonempty subset of a type can have infinitely many iterated preimages under A^{-1}. This is a rephrasing of an argument which occurred above in the discussion of the uniqueness of distinguished extensions.

Definition (extensional): We say that an element of a type is *extensional* iff it is pre-extensional and its distinguished extension has an even number of iterated images under A^{-1}. This implies that each of its other extensions has an odd number of iterated images under A^{-1}.

Properties of extensional type elements: This is enough to ensure that two extensional model elements with any common extension (over a type other than τ_{-1}) will be equal: if two extensional model elements have an empty extension (over a type other than τ_{-1}) in common, they both have all extensions empty and are equal. If two extensional model elements have a nonempty extension in common, it will be the distinguished extension of both, or a non-distinguished extension of both, since distinguished and non-distinguished extensions are taken from disjoint classes of subsets of types (when nonempty). In either case we deduce that both have the same distinguished extension and thus have all extensions the same and

are equal. Note that this gives weak extensionality over τ_{-1} (many objects have empty extension over type -1) but it gives full extensionality over any other type.

A legacy notation from earlier versions which may be used: We introduce the notation (β, δ, D) where $\delta < \beta$ and $D \subseteq \tau_\delta$. This stands for the unique extensional element x of τ_β^* such that $x \cap \tau_\delta = D$. It should be clear that there is only one such object. If D is empty, it is the unique element of τ_β^* with empty intersection with each τ_γ^* for $\gamma < \beta$. If $\delta = -1$ and D is nonempty, or if $\delta > -1$ and D has an even number of iterated images under A^{-1}, then it is the unique element x of of τ_β^* which is extensional and has distinguished extension $x \cap \tau_\delta$. If D is nonempty and has an odd number of iterated images under A^{-1}, let $A^{-1}(D) \subseteq \tau_\gamma$, and it is the same as $(\beta, \gamma, A^{-1}(D))$. This notation is mainly for compatibility with previous versions of the paper, but may have its uses.

We assume that all elements of τ_β's already constructed are extensional. This completes the mechanism for enforcement of extensionality in the structure we are defining.

3.4   Allowable permutations and supports

brief note on our further needs: A crucial aspect of this is that we will need to define τ_α so that it has cardinality μ for the process to continue. It is certainly not a sufficient restriction to require elements of τ_α to be extensional: we will require a further symmetry condition.

structural permutations defined: We define classes of permutations of our structures.

A -1-structural permutation is a permutation of τ_{-1}.

An α-structural permutation is a permutation π of τ_α such that for each type $\beta < \alpha$ there is a β-structural permutation π_β such that $\pi(x) \cap \tau_\beta = \pi_\beta{}``(x \cap \tau_\beta)$ for any $x \in \tau_\beta$.

derivatives of structural permutations: The maps π_β are referred to as derivatives of π.

More generally, for any finite subset A of $\lambda \cup \{-1\}$ with maximum α, define π_A as $(\pi_{A \setminus \{\min(A)\}})_{\min(A)}$. The maps π_A may be referred to as iterated derivatives of π. It should be clear that a structural permutation is exactly determined by its iterated derivatives which are -1-structural.

We introduce the brief notation $\pi_A^+(x) = \pi_{A \cup \{-1\}}(x)$ where $x \in \tau_{-1}$ and $-1 \notin A$ [in fact, $\pi_A^+ = (\pi_A)_{-1}$, but we find the short notation useful.]

allowable permutations defined: Structural permutations are defined on the supertype structure generally. We need a subclass of structural permutations which respects our extensionality requirements.

A -1-allowable permutation is a permutation π of τ_{-1} such that for any near-litter N, $\pi``N$ is a near-litter.

An α-allowable permutation is an α-structural permutation, each of whose derivatives π_β is a β-allowable permutation (and satisfies the condition that $\pi_\beta``\tau_\beta = \tau_\beta$) and which satisfies a coherence condition relating the f maps and derivatives of the permutation: for suitable $\beta, \gamma < \alpha$,

$$f_{\beta,\gamma}(\pi_\beta(x), \pi_\beta[S]) \sim \pi_\gamma^+ ``f_{\beta,\gamma}(x, S).$$

(where the action of allowable permutations on supports will be defined shortly).

Note that an α-allowable permutation is actually defined on the entire supertype structure, though what interests us about it is its actions on objects in our purported TTT model.

Of course, for each $-1 < \beta < \alpha$ we have already defined β-allowable permutations in the same way earlier in the recursive construction.

Definition (support condition): A β-support condition is defined as a pair (x, A), where

1. A is a finite subset of λ with maximum β (note that it will not contain -1),

2. and $x \subseteq \tau_{-1}$ is either a singleton or a near-litter.

Definition (support): Where $0 \le \beta \le \alpha$, a β-*support* is defined as a function S from a small ordinal to β-support conditions.

We may write S_δ intead of $S(\delta)$.

For any support condition (x, B) we define $(x, B)^{\uparrow A}$ as $(x, B \cup A)$ if all elements of the set A dominate all elements of the set B. Further, if S is a support, we define $S^{\uparrow A}$ so that $(S^{\uparrow A})_\epsilon = (S_\epsilon)^{\uparrow A}$. By an abuse of notation we may write $(x, B)^{\uparrow \beta}$ or $S^{\uparrow \beta}$ where β is an ordinal for $(x, B)^{\uparrow \{\beta\}})$ or $S^{\uparrow \{\beta\}}$.

For any supports S and T we denote by $S + T$ a support which consists of S, followed by T: what this means is that $(S + T)_\epsilon = S(\epsilon)$ [which we write S_ϵ] for ϵ in the domain of S, $(S + T)_{\mathrm{dom}(S)+\epsilon} = T_\epsilon$ for ϵ in the domain of T.

By fiat, we state that there is one -1-support, the empty set. Supports of atoms are never used in the argument, and it is most convenient to have a single dummy object to fill the support slot in elements of τ_{-1}^+. The Lean formalization uses a different approach but similarly gives each atom just one support, which is never used for anything.

We define the action of a β-allowable permutation π on a β-support S: if $S(\delta) = (x, A)$, $\pi[S](\delta) = (\pi_A^+``x, A)$. In the case of -1-supports S, $\pi[\emptyset] = \emptyset$.

An element x of τ_β^* ($\beta \geq 0$) has β-support S iff for every β-allowable permutation π, if $\pi[S] = S$ then $\pi(x) = x$. We say that an element x of τ_β^* which has a β-support is β-symmetric. Every element of τ_{-1} has the empty set as support (by fiat, as noted).

It is straightforward to observe that there are μ supports, since there are μ atoms, μ near-litters, and $< \kappa$ finite sets of elements of λ. Thus τ_β^+ is already known to be of size μ for $\beta < \alpha$.

It is important to note that if S is a support of $x \in \tau_\beta$, $\pi[S]$ is a support of $\pi(x)$ for any β-allowable permutation π.

motivation of the coherence condition: The motivation for this is that we need α-allowable permutations to send extensional elements of supertypes to extensional elements. Suppose $x \cap \tau_\beta = \{b\}$. If x is extensional, this has to be the distinguished extension of x. For any $\gamma \in \alpha \setminus \{\beta\}$, it follows that $x \cap \tau_\gamma$ is the set of all N_γ such that $N \sim f_{\beta,\gamma}(b, S)$ for some support S of b. This tells us that an α-allowable permutation π, for which we must have that $\pi(x)$ has β-extension $\{\pi_\beta(b)\}$, must have the γ-extension of $\pi(x)$ equal to $\pi_\gamma``\{N_\delta : \exists S : N \sim f_{\beta,\gamma}(b, S)\}$ but must also have its γ-extension equal to $\{N_\delta : \exists S : N \sim f_{\beta,\gamma}(\pi_\beta(b), S)\}$. This tells us that $\pi_\gamma(f_{\beta,\gamma}(b, S)_\delta) \in \{N_\delta : (\exists T : N \sim f_{\beta,\gamma}(\pi_\beta(b), T)\}$ for each support S of b. The coherence condition enforces this neatly, showing that it is motivated by considerations required to get extensionality to work: the action of π_β conveniently correlates supports of b with supports of $\pi_\beta(b)$.

allowable permutations preserve extensionality :

We defined $A_\delta(a)$ as

$$\{N_\delta : (\exists x \in a : (\exists S : N \sim f_{\gamma,\delta}(a)))\}.$$

If π is allowable of suitable index, $\pi_\delta``A_\delta(a) = A_\delta(\pi_\gamma``a)$ follows from the coherence condition. Verify this:

Suppose we have N_δ with $x \in a$ such that $N \sim f_{\gamma,\delta}(x, S)$. Then

$$\pi_\delta(N_\delta) \cap \tau_{-1} = (\pi_\delta)_{-1}``N \sim (\pi_\delta)_{-1}``f_{\gamma,\delta}(x, S) \sim f_{\gamma,\delta}(\pi_\gamma(x), \pi_\gamma[S]).$$

So any element of $\pi_\delta``A_\delta(a)$ is in $A_\delta(\pi_\gamma``a)$.

Suppose we have N_δ with $x \in a$ such that $N \sim f_{\gamma,\delta}(\pi_\gamma(x), S)$. We then have $N \sim (\pi_\delta)_{-1}``f_{\gamma,\delta}(x, \pi_\gamma^{-1}[S])$. We want to show that $\pi_\delta^{-1}(N_\delta) \in A_\delta(a)$. $\pi_\delta^{-1}(N_\delta) \cap \tau_{-1} = (\pi_\delta)_{-1}^{-1}``N \sim (\pi_\delta)_{-1}^{-1}``((\pi_\delta)_{-1}``f_{\gamma,\delta}(x, \pi_\gamma^{-1}[S])) = f_{\gamma,\delta}(x, \pi_\gamma^{-1}[S])$, establishing what we need.

Notice that this shows that the coherence condition implies that the image under an allowable permutation of a pre-extensional element of our structure is pre-extensional.

Now this implies that if $a \subseteq \tau_\gamma$, then $A^{-1}(a)$ exists and is in τ_δ exactly if $A^{-1}(\pi_\gamma``a)$ exists and is in $\tau_\delta$, and moreover $A^{-1}(\pi_\gamma``a)$ is equal

to π_δ "$A^{-1}(a)$ if it exists under these conditions. This verifies that the coherence condition implies that allowable permutations preserve full extensionality, not just pre-extensionality: the number of iterated images under A^{-1} of an extension that exist is not affected by application of an allowable permutation in a suitable sense.

the definition of τ_α: We stipulate that all elements of τ_β have β-supports, and define τ_α as the set of elements x of τ_α^* such that $x \cap \tau_\beta^* \subseteq \tau_\beta$ for each $\beta < \alpha$, x is extensional, and x has an α-support.

We still have to prove that the cardinality of τ_α, and so of τ_α^+, is μ, to show that the construction works.

Observation (κ-completeness of the structure): For any subset X with cardinality $< \kappa$ of a type γ and $\beta > \gamma$, it should be clear that X_β has a support, whose range is obtained from the union of the ranges of the supports of the elements of X by replacing each element (u, B) of the union of the ranges with $(u, B \cup \{\beta\})$, and therefore belongs to the model. X_β is obviously extensional (the extension X is clearly the distinguished extension and has no image under A^{-1}).

conditions on choice of distinguished ordinal indexings of supported types: $\iota_*^+(x, \emptyset)$ is defined as $\iota_*(x)$. The well-ordering $\leq_{-1}^+$ is defined by

$$(x, \emptyset) \leq_{-1}^+ (y, \emptyset) \leftrightarrow \iota_*(x) < \iota_*(y).$$

The conditions constraining choice of functions ι_β^+ $(-1 < \beta \leq \alpha)$ are

1. $\iota_*(t) < \iota_*^+(x, S)$ if (t, A) is in the range of S and x is not a typed atom or near-litter

2. $\iota_*(t) \leq \iota_*^+(t_\gamma, S)$

We then define $x \leq_\beta^+ y$ as $\iota_\beta^+(x) \leq \iota_\beta^+(y)$.

There is no difficulty in satisfying these constraints as there are only a small set of constraints on any particular value of ι_*^+ and μ is of cofinality at least κ.

Of course, if $\beta < \alpha$ this records a hypothesis of the recursion: one of these is noted above already.

It should be noted that type 0 has a very simple description: the -1-extensions of type 0 objects are exactly the sets with small symmetric difference from small or co-small unions of litters, and that these are the same extensions over type -1 which appear in any positive type.

At this point we have a complete description of the structure which we claim is a model of TTT.

4 Verification that the structure defined is a model

4.1 The freedom of action theorem

Definition (approximation): A β-approximation is a map π^0 from finite subsets of λ with maximum element β such that each $\pi^0(A)$ (which we write π_A^0) is a function with the following properties:

1. The domain and image of π_A^0 are the same and π_A^0 is injective.

2. Each domain element x of π_A^0 is such that (x, A) is a support condition.

3. x and $\pi_A^0(x)$ have the same cardinality, which is either 1 or κ, since the previous condition tells us that x is a singleton of an atom or a near-litter.

4. if the cardinality of x is κ, x and $\pi^0(x)$ are litters.

5. for any x in the domain of π_A^0, the collection of y in the domain of π_A^0 such that $y \subseteq x$ is small.

We say that π^0 approximates a β-allowable permutation π just in case $\pi_A^+ \text{``} x = \pi_A^0(x)$ if $x$ has cardinality 1, and $\pi_A^+ \text{``} x \sim \pi_A^0(x)$ otherwise, for each (x, A) in the domain of π^0.

Notice that each such π^0 has an inverse $(\pi_0)^{-1}$ determined by $(\pi^0)_A^{-1} = (\pi_A^0)^{-1}$, which is also a β-approximation.

litter near a near-litter: Recall that any near-litter N, we define N° as the unique litter L such that $L \sim N$.

Definition (flexibility): A near-litter x is A-flexible if x° is not in the range of any $f_{\gamma, \min(A)}$ for $\gamma < \min(A_1)$.

Definition (exception, exact approximation): A -1-allowable permutation π has *exception* x if, L being the litter containing x, we have either $\pi(x) \notin (\pi \text{``} L)^\circ$ or $\pi^{-1}(x) \notin (\pi^{-1} \text{``} L)^\circ$.

A β-approximation π^0 *exactly approximates* a β-allowable permutation π iff π^0 approximates π and for every exception x of a $\pi_{A \cup \{-1\}}$ (A not containing -1) we have $\{x\}$ in the domain of π_A^0.

Theorem (freedom of action): A β-approximation π^0 will exactly approximate some β-allowable permutation π if it satisfies the additional condition that any domain element x of π_A^0 which is a litter is A-flexible.

Proof: For each pair of sets L, M which are co-small subsets of litters, we define $\pi_{L,M}$ as the unique map ρ from L onto M such that for any $x, y \in L$,

$$x <_{L^\circ} y \leftrightarrow \rho(x) <_{M^\circ} \rho(y):$$

$\pi_{L,M}$ is the unique map from L onto M which is strictly increasing in the order determined by fourth projections of atoms. Notice that $\pi_{M,N} \circ \pi_{L,M} = \pi_{L,N}$ will hold if L, M, N are subsets of the same litter, and $\pi_{L,M} \circ \pi_{M,L} = \pi_{L,L}$ which is the identity map on L, under the same conditions. The relationship to composition is neatly handled by our concrete definition.[2]

We also choose an extension of each π_A^0 to all A-flexible litters; we do this without notational comment, simply assuming that π_A^0 is defined for each A at each A-flexible litter M, which can be arranged harmlessly (for example, one could have π_A^0 act as the identity on the new A-flexible litters, but we do not require this).

We choose an approximation π^0 satisfying the conditions of the theorem and extend it as indicated in the previous paragraph. We compute the allowable permutation π, and in parallel its inverse π^{-1}, on all support conditions (and therefore compute all its derivatives π_A (and π_A^{-1}) at all atoms ($-1 \in A$), so completely defining it, using the assumption that we already know how to carry out this construction to define γ-allowable permutations exactly approximated by any given γ-approximation for $\gamma < \beta$.

The basic order of the construction of π is that the value of $(\pi_A^+)^i {}^{\shortmid\shortmid}x$ for $i = \pm 1$ is computed at each atom or near-litter x assuming that the values have already been computed for $(\pi_B^+)^j {}^{\shortmid\shortmid}y$ for each $i = \pm 1$, y with $\iota_*(y) < \iota_*(x)$, and extended type index B [of course, this is equivalent to computing $\pi_A(x_{\mathtt{min}(A)})$ and we may discuss it in that form]. All specific discussions of inductive hypotheses below conform to this.

We use the notation π_A^* for the partially computed π_A at any point in the calculation before we are done. We use the notation $\pi_A^{+*}(x)$ for $\pi_{A \cup \{-1\}}^*(x)$. We use similar notation $(\pi_A^{-1})^*$ for the part of π^{-1} which we have already computed.

We first indicate how to compute $\pi_A^*(L_{\mathtt{min}(A)})$, where L is a litter. We compute $(\pi_A^{-1})^*(L_{\mathtt{min}(A)})$ in the same way.

We further extend this, once $\pi^*(L_{\mathtt{min}(A)})^\circ$ is computed as $M_{\mathtt{min}(A)}$, to describe the action of π_A^{+*} on elements of L: for each $x \in L$, if $\{x\}$ is in the domain of π_A^0, which maps it to $\{y\}$, π_A^{+*} maps x to y. Define L^- as the set of all $x \in L$ such that $\{x\}$ is not in the domain of π_A^0 and define M^- as the set of all $x \in M$ such that $\{x\}$ is not in the domain of π_A^0. For $x \in L^-$ we define $\pi_A^{+*}(x)$ as $\pi_{L^-,M^-}(x)$. Note that this approach ensures that, where π is the permutation we eventually construct, there can be no exceptions of π_A^+ other than sole elements of elements of the domain of π_A^0.

---

[2]the choice of these maps does not need to be so concrete, but the fact that it can be indicates for example that there is no use of choice here. We like the concreteness of this approach.

We can then exactly compute $\pi_A^*(L_{\min(A)})$ as $(\pi_A^{+*}\text{``}L)_{\min(A)}$

Note that the general inductive hypothesis ensures that when we are computing $\pi_A^+(x)$ and L is the litter containing x, we have already computed $\pi_A^+\text{``}L$, because $\iota_*(L) < \iota_*(\{x\})$. The paragraph above indicates how this computation is carried out for every atom and extended type index.

If L is A-flexible, we can compute $\pi_A^*(L_{\min(A)})^\circ$ as $(\pi_A^0(L))_{\min(A)}$

If L is A-inflexible, we have $f_{\gamma,\min(A)}(x, S) = L$ for some $\gamma < \min(A_1)$, $x \in \tau_\gamma$, and S a support of x.

We expect $\pi_A^*(f_{\gamma,\min(A)}(x, S)_{\min(A)})$ to have -1-extension near that of

$$f_{\gamma,\min(A)}(\pi_{A_1\cup\{\gamma\}}^*(x), \pi_{A_1\cup\{\gamma\}}^*[S]))_{\min(A)}.$$

As a comment on the circumlocution above, we allow ourselves subsequently to say that M_δ is near N_δ when $M \sim N$: nearness of typed near-litters is defined in the obvious way in terms of nearness of near-litters.

If $\gamma > -1$, we will show that the inductive hypotheses allow us to compute values of $\pi_{A_1\cup\{\gamma\}}^*$ and so images under this permutation of γ-supports.

The conditions on the construction of the indexings ι_β^+ ensure that we can compute $\pi_{A_1\cup\{\gamma\}}^*[S]$ and $(\pi_{A_1\cup\{\gamma\}}^{-1})^*[S]$, because $\iota_*(u) < \iota_*^+((x, S)) < \iota_*(L)$ for each $(u, U) \in S$ so we have already computed each $\pi_*^+\text{``}u$.

We intend to construct a γ-approximation ρ which must send S to $\pi_{A_1\cup\{\gamma\}}^*[S]$. Each atomic and flexible item in S is to be mapped by the γ-approximation to the corresponding item in $\pi_{A_1\cup\{\gamma\}}^*[S]$ (with the proviso that where near-litter items correspond, they will be modified to litter items). We have to fill in orbits, since each ρ_B^0 must have domain the same as its range. We fill in the orbits, adding at most a countable number of flexible litters (which present no difficulties) and a countable number of atoms to domains of ρ_B^0 per atom or flexible near-litter already there, subject to the condition that where we are choosing the image or preimage of the singleton of an atom under ρ_B^0 and we know its elementwise image or preimage under $\pi_{A_1\cup B}^{+*}$ we use that (and we will know this if we know the elementwise image or preimage under $\pi_{A_1\cup B}^{+*}$ of a litter which includes it, as in the case of litters near the near-litters in the support). Where we do not have this information, we can choose images and preimages freely as long as they are not already known elementwise preimages or images under $\pi_{A_1\cup B}^{+*}$.

There is a further point to do with anomalous elements of near-litters N with $(N, B) \in S$ which are not litters: for any $z \in N \setminus N^\circ$, with z belonging to a litter M, $\iota_*(M) < \iota_*(\{z\}) < \iota_*^+(x, S) < \iota_*(L)$ so $\pi_B^+\text{``}\{z\}$ and $\pi_B^+\text{``}M$ have both been computed, and appropriate values can be assigned to ρ^0 at $(\{z\}, B)$ and (M, B), to ensure that ρ will act correctly on N.

We construct ρ exactly approximated by ρ^0. We need to verify that $\rho[S]$ really can be relied upon to agree with $\pi_{A_1\cup\{\gamma\}}^*[S]$. The difficulty is that ρ^0

agrees with $\pi^*_{A_1\cup\{\gamma\}}$ for each support condition in S which has first component the singleton of an atom or flexible, because this information was packed into ρ^0: how do we know that it agrees with $\pi^*_{A_1\cup\{\gamma\}}$ at inflexible items? Suppose it failed to agree: there would be an ι^+_* minimal item at which disagreement occurred. If the support condition were $(N, A_1 \cup D)$, we know that ρ_D and $\pi^*_{A_1\cup D}$ agree up to nearness at N_ϵ ($\epsilon = \mathtt{min}(D)$) because the actions of ρ and $\pi^*_{A_1\cup\{\gamma\}}$ agree on the support appearing in the inverse image under the appropriate f map of N°. Further, we have ensured by construction that any exceptional action of ρ_B in N will agree with the expected action of $\pi_{A_1\cup B}$: any exceptional action of ρ_B must either agree with computed action of $\pi^*_{A_1\cup B}$ or be in a litter for which we have computed no value for $\pi_{A_1\cup B}$. Thus we get not just nearness but identity.

Now we compute $\pi^*_{A_1\cup\{\gamma\}}(x)$ as $\rho(x)$, and so we know $\pi^*_A(f_{\gamma,\mathtt{min}(A)}(x,S)_{\mathtt{min}(A)})$ up to nearness, because we know how to compute

$$f_{\gamma,\mathtt{min}(A)}(\pi^*_{A_1\cup\{\gamma\}}(x), \pi^*_{A_1\cup\{\gamma\}}[S]))_{\mathtt{min}(A)}.$$

If $\gamma = -1$, we need to compute $\pi^*_{A_1\cup\{-1\}}(x)$, for which it is sufficient to compute $\pi^*_{A_1}(M_{\mathtt{min}(A_1)})$, where M is the litter containing x. We have $\iota_*(M) < \iota^+(x,\emptyset) < \iota_*(L)$ so by inductive hypothesis stated above we have already computed $\pi^*_{A_1}(M_{\mathtt{min}(A_1)})$, so we can compute $\pi^*_{A_1}(x)$, and so we can compute $\pi^*_A(L)$ up to nearness.

In this way we have computed $\pi^*_A(L_{\mathtt{min}(A)})^\circ$ in every case and above we indicate how to compute $\pi^*_A(L_{\mathtt{min}(A)})$ exactly given this.

The process given will compute $\pi_A(x)$ and $\pi_A^{-1}(x)$ for every atom x and every A containing -1. Since every action on every atom is fixed, π is fixed as a structural permutation.

The method by which the derivatives of π are evaluated at atoms ensures that π_A agrees with π^0_A on typed singletons. It also ensures that (if π and its derivatives defined as indicated satisfy the coherence conditions) $\pi_{A\cup\{-1\}}$ has an exception x only if $\{x\}$ is in the domain of π^0_A.

The method of computation verifies that the coherence conditions will hold. The method of computation also verifies that π is a permutation, as π^{-1} is computed in precisely the same way from $(\pi^0)^{-1}$.

4.2 Types are of size μ (so the construction actually succeeds)

Now we argue that (given that everything worked out correctly already at lower types) each type α is of size μ, which ensures that the construction actually succeeds at every type.

Definition (strong support): A β-*strong support* is a β-support S with the additional properties that

1. if (x, A) and (y, A) are in the range of S and $x\Delta y$ is small, then $(\{z\}, A) \in S$ for all $z \in x\Delta y$,

2. and for each ϵ in the domain of S, if $S_\epsilon = (x, A)$, and $x^\circ = f_{\gamma,\delta}(y, T)$, then the range of $T^{\uparrow A_1}$ is a subset of the range of $S\lceil\epsilon$: supports appearing in inverse images under f of litters which are near the first projections of an element of the support have a type-raised copy (mod reindexing) appearing in the support before that item.

Remark: It should be evident that if π is a β-allowable permutation and S is a β-strong support, $\pi[S]$ is also a β-strong support.

Remark: Each support S is the terminal segment of a strong support. Such a strong support can be constructed by prefixing to S, for each $T^{\uparrow A_1}$ such that for some ϵ in the domain of S, $S_\epsilon = (x, A)$, and $x^\circ = f_{\gamma,\delta}(y, T)$, a strong support with $T^{\uparrow A_1}$ as a terminal segment, which will be obtainable as $U^{\uparrow A_1}$ where U is a strong support with T as a terminal segment, which exists by the inductive hypothesis that this is true for supports with lower type index than S, followed by all atomic items which must be added to satisfy the first condition in the definition of strong support. In fact, we can define a canonical strong support of which S is a terminal segment by stipulating that the $U^{\uparrow A_1}$'s are added (each one in turn between the preceding ones and S) in the order in which the correlated S_ϵ's appear in S and that U is the canonical downward extension of T in each case, and that the atomic items are added in the order of their images under ι_*^+, with some fixed well-ordering of extended type indices used to resolve order of items with the same value under ι_*^+.

Definition (coding functions): For any support S and object x, we can define a function $\chi_{x,S}$ which sends $T = \pi[S]$ to $\pi(x)$ for every T in the orbit of S under the action of allowable permutations. We call such functions *coding functions*. Note that if $\pi[S] = \pi'[S]$ then $(\pi^{-1} \circ \pi')[S] = S$, so $(\pi^{-1} \circ \pi')(x) = x$, so $\pi(x) = \pi'(x)$, ensuring that the map $\chi_{x,S}$ for which we gave an implicit definition is well defined.

designated supports: For each ordinal γ, and for each orbit in τ_γ under allowable permutations, choose x in the orbit (the designated element of the orbit), choose a strong support S of x, and for each γ-allowable permutation π define the designated support of $\pi(x)$ as $\pi[S]$.

Definition (specification): A χ-specification S^* of a χ-support S is a function with the same domain as S. We use the notation S_ϵ^* for $S^*(\epsilon)$.

1. If S_ϵ is $(\{x\}, A)$, where $\beta = \mathtt{min}(A)$, then S_ϵ^* is $(0, \beta, \Sigma, A)$ where Σ is the set of all δ such that $A = \pi_2(S_\delta)$ and $x \in \pi_1(S_\delta)$ (this captures identical atoms and near litters containing the given atom)

2. If S_ϵ is (N, A) and N is a near-litter, where $\beta = \mathtt{min}(A)$, and either $|A| = 1$ or N° is not in the range of any $f_{\gamma,\beta}$ for $\gamma < \mathtt{min}(A_1)$, then S_ϵ^* is $(1, \beta, \Sigma, A)$, where Σ is the set of all δ such that $\pi_2(S_\delta) = A$ and $\pi_1(S_\delta) \sim \pi_1(S_\epsilon)$.

3. If S_ϵ is (N, A) and N is a near-litter, where $\beta = \mathtt{min}(A)$, and $N^\circ = f_{\gamma,\beta}(x, T)$ with $-1 < \gamma < \mathtt{min}(A_1)$ and $x \in \tau_\gamma$ then S_ϵ^* is $(2, \beta, \chi_{x,T}, F, A)$, where F is a function from the domain of T into ϵ such that $S_{F(\delta)} = (T^{\uparrow A_1})_\delta$ for each δ in the domain of T, or 1 if there is no such F (the usefulness of 1 as a dummy being that it is not a function). There is a method of choose a canonical such F if there is one: add the provision that for each δ, $F(\delta)$ is chosen as small as possible.

4. If S_ϵ is (N, A) and N is a near-litter, where $\beta = \mathtt{min}(A)$, and $N^\circ = f_{-1,\beta}(x, \emptyset)$ then S_ϵ^* is $(3, \beta, \Sigma, A)$, where Σ is the set of all δ such that S_δ is $(\{x\}, A_1)$.

Remark: It should be evident that every support has a specification, and that a strong support will have a specification with no instances of $F = 1$, and that for any β-allowable permutation π and strong β-support S, $(\pi[S])^* = S^*$. What is less evident and our first target result here is that if S is a strong support then any T with $T^* = S^*$ is the image of S under the action of an allowable permutation: the specifications precisely code the orbits in the strong supports under the allowable permutations.

Observation: On the inductive hypothesis that there are $< \mu$ γ-coding functions and $< \mu$ γ-specifications for each $\gamma < \beta$, we observe that there are $< \mu$ specifications of β-supports for $\beta \leq \alpha$.

Lemma: The specification(s) of a strong β-support exactly determine the orbit in the action of β-allowable permutations on supports to which it belongs: if two β-supports have the same specification, they are in the same orbit.

Proof of Lemma: It is straightforward to see that if S is a β-support and if π is a β-allowable permutation, and S^* is the specification for S, that S^* is also the specification for $\pi[S]$. The relationships between items in the support recorded in the specification are invariant under application of allowable permutations.

It remains to show that if S and T are supports, and $S^* = T^*$ is a specification for both, there is an allowable permutation π such that $\pi[S] = T$.

We construct π using the Freedom of Action Theorem.

If we have $S_\epsilon = (\{x\}, A)$, we will have $T_\epsilon = (\{y\}, A)$ for some y, and we will set $\pi_A^0(\{x\}) = \{y\}$ as part of the construction of the local bijection to be used. Note that if $S_\delta = (\{x\}, A)$ for $\delta \neq \epsilon$, the fact that S and T have the same specification ensures that $T_\delta = (\{y\}, A)$, because the specification contains the information that the specified supports have the same values at ϵ and δ.

We suppose all these approximation values are computed at the outset. Further, we fill in orbits, with the proviso that if an atom x to be sent by π_A^0 to an atom y and (M, A) is in S with the same index as (N, A) in T, then $x \in M$ if and only if $y \in N$, and if an atom x to be sent by $(\pi_A^0)^{-1}$ to an atom y and (M, A) is in T with the same index as (N, A) in S, then $x \in M$ if and only if $y \in N$. There is no obstruction to choosing values to meet these conditions, sufficiently to fill orbits.

If we have $S_\epsilon = (M, A)$ for M a near litter and either $|A| = 1$ or M° is not in the range of any $f_{\gamma,\beta}$ for $\gamma < \min(A_1)$, then $T_\epsilon = (N, A)$ for N a near litter, with analogous properties, and we set $\pi_A^0(M^\circ) = N^\circ$ as part of the data for application of the Freedom of Action Theorem. Note again that information in the specification ensures that if $S_\delta = (M', A)$ for $M' \sim M$ that $T_\delta = (N', A)$ will have $N' \sim N$ because the specification tells us that T has first components of values being litters near one another at the same pairs of ordinals that S does, and so no conflicting computation of values for the approximation will occur.

We suppose orbits filled in in the flexible litters, which can be done quite freely.

In the computation for the case of inflexible litters, we suppose that the computation of all orbits under the approximation has been completed for all atomic and flexible items in S, and for inflexible items appearing earlier in the support.

If we have $S_\epsilon = (M, A)$ for M a near litter with $M^\circ = f_{\gamma,\beta}(x, U)$, where $-1 < \gamma < \min(A_1)$, then $S_\epsilon^* = T_\epsilon^*$ is $(2, \beta, \chi_{x,U}, F, A)$ and where $T_\epsilon^+ = (N, A)$, we have $N^\circ = f_{\gamma,\beta}(y, V)$ where $\chi_{y,V} = \chi_{x,U}$, so any allowable permutation π such that $\pi_{A_1 \cup \{\gamma\}}[S \circ F] = T \circ F$ will satisfy $\pi_{A_1 \cup \{\gamma\}}(x) = y$ and so $\pi_A^+ \text{``}M \sim N$. We add additional information to our approximation to make it so M will be sent precisely to N. Extend the approximation so that π_A^0 sends each singleton of an element of $M \setminus M^\circ$ to a singleton of an element of N° and $(\pi_A^0)^{-1}$ sends each singleton of an element of $N \setminus N^\circ$ to a singleton of an element of M°. One needs to further extend the approximation to fill in orbits. The constraint must be observed in making extensions at atoms that if P is a litter with a value for $\pi_A^0(P)$ already determined, to which x belongs [resp. does not belong] then $\pi_A^0(\{x\})$ must be chosen as $\{y\}$ such that y belongs to [does not belong to] $\pi_A^0(P)$ [and there is a precisely parallel condition for choosing new

41

values of $(\pi_A^0)^{-1}$ to fill orbits]. This strategy avoids introduction of problematic exceptions. We then add $\pi_A^0(M^\circ) = N^\circ$ to the approximation for bookkeeping purposes: we have arranged for the atomic and flexible items in the approximation to force this valuation anyway [so the applicability of Freedom of Action is not affected; inflexible items can be dropped from the approximation before the theorem is applied], and we need this information to guide orbit filling. Note that there is enough computational information in the specification to fill in orbits in the approximation for inflexible litters too (which will require the same adjustment for exact fit at each step).

If we have $S_\epsilon = (M, A)$ for M a near litter with $M^\circ = f_{-1,\beta}(x, \emptyset)$, then we have $T_\epsilon = (N, A)$ where $N^\circ = f_{-1,\beta}(y, \emptyset)$, and we add to our approximation the information that $\{x\}$ is mapped to $\{y\}$ by $\pi_{A_1}^0$ (and fill in orbits), and the fix to extend the approximation to get M to map precisely to N elementwise under π_A^+ is exactly as in the previous case. The usual observation can be made that if (M, A) appears at another location in S, there is enough information in the specification to force the computation to give the same result.

If we have $S_\epsilon = (M, A)$ for M a near-litter and $\delta < \epsilon$ such that $S_\delta = (M', A)$ with $M \sim M'$, then observe that for each z in $M \Delta M'$, $(z, A) \in S$. The computations already done to force S_ϵ to map to T_ϵ, along with the easy calculations for atoms, show that there is no additional work needed in this case, and so we do not do any, and no conflict arises.

In the indicated way, we construct an approximation such that the allowable permutation it exactly approximates must have action sending S to T, completing the proof of the Lemma.

There are not too many orbits: Since the β-specifications $[\beta \leq \alpha]$ precisely determine the orbits in supports under β-allowable permutations, and there are $< \mu$ β-specifications (on stated hypotheses) there are $< \mu$ such orbits.

weak specifications: Notice that we can give a kind of specification for any support S: give the specification for a strong support T of which S is an end extension and the index at which S starts: this will determine the orbit in which S lies in the allowable permutations. This establishes that the collection of orbits in the β-supports is no larger than the collection of orbits in the strong β-supports. These weak specifications are not unique.

The strategy of our argument for the size of the types is to show that that there are $< \mu$ coding functions for each type , which implies that there are no more than μ (and so exactly μ) elements of each type, since every element of a type is obtainable by applying a coding function (of which there are $< \mu$) to a support (of which there are μ).

Analysis of coding functions for type 0: We describe all coding functions for type 0. The orbit of a 0-support in the allowable permutations is

determined by the positions in the support occupied by near-litters, and for each position in the support occupied by a singleton, the position, if any, of the near-litter in the support which includes it. There are no more than 2^κ ways to specify an orbit. Now for each such equivalence class, there is a natural partition of type -1 into near-litters, singletons, and a large complement set. The partition has $\nu < \kappa$ elements, and there will be $2^\nu \leq 2^\kappa$ coding functions for that orbit in the supports, determined by specifying for each compartment in the partition whether it is to be included or excluded from the set computed from a support in that orbit. So there are no more than $2^\kappa < \mu$ coding functions over type 0.

Analysis of the general case:

Our inductive hypothesis is that for each $\beta < \alpha$ we have $< \mu$ β-coding functions.

We specify an object $X \in \tau_\alpha$ and a strong α-support S for X, and develop a recipe for the coding function $\chi_{X,S}$ which can be used to see that there are $< \mu$ α-coding functions (assuming of course that we know that things worked out correctly for $\beta < \alpha$).

$X = B_\alpha$, where B is a subset of τ_β.

We define S_b as the designated strong support for b, and T_b as the canonical extension to a strong support of $S_b^{\uparrow\alpha} + S$.

For each $b \in B$ there is a support S_b chosen as above, from which the support T_b can be computed as described above. If $b' \in B$ is in the range of the same coding function χ_{b,S_b} as b, $S_{b'}$ is $\pi[S_b]$ for some β-allowable π with $\pi(b) = b'$. If we have the further condition that T_b and $T_{b'}$ have the same specification, it follows that there is a permutation π_2 such that $\pi_2[T_b] = T_{b'}$. Note that $(\pi_2)_\beta[S_b] = S_{b'}$, from which it follows that $\pi^{-1} \circ (\pi_2)_\beta$ fixes b, since it fixes all elements of S_b, so $b' = \pi(b) = (\pi_2)_\beta(b)$, from which it follows that $\pi_2(\{b\}_\alpha) = \{b'\}_\alpha$ so $\{b\}_\alpha$ and $\{b'\}_\alpha$ are in the range of the same coding function $\chi_{\{b\}_\alpha,T_b}$. Now there are $< \mu$ possible specifications of a coding function χ_{b,S_b} [we know that there are $< \mu$ βcoding functions] followed by a specification for T_b [we know that there are $< \mu$ α-specifications], so by this procedure we describe a family of $< \mu$ coding functions $\chi_{\{b\}_\alpha,T_b}$ whose range covers all type α singletons of elements of B.

We claim that $\chi_{X,S}$ can be defined in terms of the orbit of S in the allowable permutations and the set of coding functions $\chi_{\{b\}_\alpha,T_b}$ for $b \in B$. There are $< \mu$ coding functions of this kind [using the fact that a set of size $< \mu$ (in this case the set of all coding functions $\chi_{\{b\}_\alpha,T_b}$ for $b \in \tau_\beta$) has $< \mu$ subsets], and we have shown above that there are $< \mu$ orbits in the α-strong supports under allowable permutations, so this will imply that there are $\leq \mu$ elements of type α (it is obvious that there are $\geq \mu$ elements of each type). Of course we get $\leq \mu$ codes for each $\beta < \alpha$, but we know that $\lambda < \kappa < \mu$.

The definition that we claim works is that $\chi_{X,S}(U) = B'_\alpha$, where B' is the set of all $\bigcup(\chi_{\{b\}_\alpha, T_b}(U') \cap \pi_\beta)$ for $b \in B$ and U a terminal segment of U'. Clearly this definition depends only on the orbit of S and the set of coding functions T_b derived from B as described above. Before we know that this is actually the coding function desired, we will write it as $\chi^*_{X,S}$.

The function we have defined is certainly a coding function, in the sense that $\chi^*_{X,S}(\pi[S]) = \pi(\chi^*_{X,S}(S))$. What requires work is to show that $\chi^*_{X,S}(S) = X$, from which it follows that it is in fact the intended function.

Clearly each $b \in B$ belongs to $\chi^*_{X,S}(S)$ as defined, because $b = \bigcup(\chi_{\{b\}_\alpha, T_b}(T_b) \cap \tau_\beta)$, and T_b has S as a terminal segment.

An arbitrary $c \in \chi^*_{X,S}(S)$ is of the form $\bigcup(\chi_{\{b\}_\alpha, T_b}(U) \cap \tau_\beta)$, where U has S as a terminal segment and of course must be in the orbit of T_b under allowable permutations, so some $\pi_0[T_b] = U$. Now observe that $\pi_0[S] = S$, so $\pi_0(X) = X$, so $(\pi_0)_\beta "B = B$. Further $(\pi_0)_\beta(b) = c$, so in fact $c \in B$ which completes the argument. The assertion $(\pi_0)_\beta(b) = c$ might be thought to require verification: the thing to observe is that $c = \bigcup(\chi_{\{b\}_\alpha, T_b}(U) \cap \tau_\beta) = \bigcup(\pi_0(\chi_{\{b\}_\alpha, T_b}(S) \cap \tau_\beta) = \bigcup(\pi_0(\{b\}_\alpha) \cap \tau_\beta) = \bigcup(\{(\pi_0)_\beta(b)\}_\beta \cap \tau_\beta) = (\pi_0)_\beta(b)$

This completes the proof: any element of a type is determined by a support (of which there are μ) and a coding function (there are $< \mu$ of these, so a type has no more than μ elements (and obviously has at least μ elements).

4.3 The structure is a model of predicative TTT

There is then a very direct proof that the structure presented is a model of predicative TTT (in which the definition of a set at a particular type may not mention any higher type). Use E for the membership relation $\in_{TTT}$ of the structure defined above (in which the membership of type β objects in type α objects is actually a subrelation of the membership relation of the metatheory, a fact inherited from the scheme of supertypes). It should be evident that $xEy \leftrightarrow \pi_\beta(x)E\pi(y)$, where x is of type β, y is of type α, and π is an α-allowable permutation.

Suppose that we are considering the existence of $\{x : \phi^s\}$, where ϕ is a formula of the language of TST with $\in$ translated as E, and s is a strictly increasing sequence of types. The truth value of each subformula of ϕ will be preserved if we replace each u of type $s(i)$ with $\pi_{A_{s,i}}(u)$, where $A_{s,i}$ is the set of all s_k for $i \leq k \leq j+1$ [x being of type $s(j)$, and there being no variables of type higher than $s(j+1)$]: $\pi_{A_{s,i}}(x)E\pi_{A_{s,i+1}}(y)$ is equivalent to $(\pi_{A_{s,i+1}})_{s(i)}(x)E\pi_{A_{s,i+1}}(y)$, which is equivalent to xEy by the observation above. The formula ϕ will contain various parameters a_i of types $s(n_i)$ and it is then evident that the set $\{x : \phi^s\}$ will be fixed by any $s(j+1)$-allowable permutation π such that $\pi_{A_{s,n_i}}$ fixes a_i for each i. But this means that $(s(j+1), s(j), \{x : \phi^s\})$ is symmetric and belongs to type $s(j+1)$: we can merge the supports of the a_i's (with suitable raising of indices) into a single $s(j+1)$-support. Notice that we assumed the predicativity condition that no variable more than one type higher than x appears (in the sense of TST).

This procedure will certainly work if the set definition is predicative (all bound variables are of type no higher than that of x, parameters at the type of the set being defined are allowed), but it also works for some impredicative set definitions.

There are easier proofs of the consistency of predicative tangled type theory; there is a reason of course that we have pursued this one.

It should be noted that the construction given here is in a sense a Frankel-Mostowski construction, though we have no real need to reference the usual FM constructions in ZFA here. Constructions analogous to Frankel-Mostowski constructions can be carried out in TST using permutations of type 0; here we are doing something much more complicated involving many permutations of type -1 which intermesh in precisely the right way. Our explanation of our technique is self-contained, but we do acknowledge this intellectual debt.

4.4 Impredicativity: verifying the axiom of union

What remains to complete the proof is that typed versions of the axiom of set union hold. That this is sufficient is a fact about predicative type theory. If we have predicative comprehension and union, we note that for any formula ϕ, $\{\iota^k(x) : \phi(x)\}$ will be predicative if k is taken to be large enough, then application of union k times to this set will give $\{x : \phi(x)\}$. $\iota(x)$ here denotes $\{x\}$. It is evidently sufficient to prove that unions of sets of singletons exist.

So what we need to show is that if $\alpha > \beta > \gamma$ and $G \subseteq \tau_\gamma$, and

$$\{\{g\}_\beta : g \in G\}_\alpha$$

is symmetric (has an α-support, so belongs to τ_α), then G_β is symmetric (has a β-support, so belongs to τ_β).

Suppose that $\{\{g\}_\beta : g \in G\}_\alpha$ is symmetric. It then has a support S. We claim that $S_{(\beta)}$, defined as $\{(z, C) : \mathtt{max}(C) = \beta \wedge (z, C \cup \{\alpha\}) \in S$, is a β-support for G_β.

Any $g \in G$ has a γ-support T which extends $(S_{(\beta)})_{(\gamma)}$.

Suppose that the action of the β-allowable permutation π fixes $S_{(\beta)}$.

Our plan is to use freedom of action technology to construct a permutation π^* whose action on S is the identity and whose action on $T^{\{\alpha,\beta\}}$ precisely parallels the action of π on $T^{\{\beta\}}$.

If this is accomplished, then the action of π^* fixes S and so fixes

$$\{\{g\}_\beta : g \in G\}_\alpha,$$

while at the same time $(\pi_\beta^*)_\gamma$ agrees with π_γ on G. This implies that $\pi_\gamma(g) \in G$ (and the same argument applies to π^{-1}) so π fixes $\{\{g\}_\beta : g \in G\}$.

Close up the γ-support T under the processes of action of π and inclusion of atoms at which π acts exceptionally to obtain T^*.

We construct the allowable permutation π^* by Freedom of Action so that the action of $(\pi_\beta^*)_\gamma$ on atomic and flexible items in T^* agrees with the action of π_γ on T^* and the action of π_* fixes atomic and flexible items in S. On any non-flexible litter L in S, π_β^* acts correctly because it acts correctly on a support of the inverse image of L under the appropriate f map (fixing all of its elements), and there will be no unexpected exceptional actions because the permutation is constructed by Freedom of Action, so we get identity rather than nearness. The tricky case seems to require a little extra attention to the action on T^*: if a non-flexible litter has inverse image u under $f_{-1,\gamma}$, it is mapped by π to something with inverse image v under $f_{-1,\gamma}$, we arrange for the approximation generating π^* to induce π_β^* to map $\{u\}_\beta$ to $\{v\}_\beta$. Thus $(\pi_\beta^*)_\gamma$ maps g to $\pi_\gamma(g)$ as required for the argument above. That said, any non-flexible item is sent to its image under the appropriate derivative of π because a support is acted on correctly and there will be no exceptional actions of derivatives of π^* disagreeing with exceptional actions of π because T^* is closed under exceptional actions of π in litters. This completes the argument.

Note that difficult interactions with S are avoided because an incompatibility of π with fixing S would involve moving most elements of a litter in the range

of $f_{-1,\beta}$, and while π may do this, nothing in the definition of π^* can force this to happen; there is no conflict between the conditions imposed by S and the conditions imposed by T^*.

This completes the proof. In the formal proof in Lean, what is actually done is a proof that each of the assertions in the finite axiomatization of Hailperin in the version discussed in the final subsection of section 2 holds in all typed versions in our structure for the language of TTT, so it is in fact a model of TTT. The axioms in Hailperin other than the axiom of type lowering are predicative comprehension axioms and admit demonstration by the methods of the previous subsection, done explicitly without metamathematics. In the formalization, the axiom of type lowering, which contains rather more content than the axiom of set union restricted to sets of singletons which is proved here, is proved by first proving the existence of an iterated image under elementwise application of the singleton operation of the desired set, whose definition is predicative, then repeatedly applying the result of this section that sets of singletons have unions.

5 Conclusions, extended results, and questions

This is a rather boring resolution of the NF consistency problem.

NF has no locally interesting combinatorial consequences. Any stratified fact about sets of a bounded standard size which holds in ZFC will continue to hold in models constructed using this strategy with the parameter κ chosen large enough. That the continuum can be well-ordered or that the axiom of dependent choices can hold, for example, can readily be arranged. Any theorem about familiar objects such as real numbers which holds in ZFC can be relied upon to hold in our models (even if it requires Choice to prove), and any situation which is possible for familiar objects is possible in models of *NF*: for example, the Continuum Hypothesis can be true or false. It cannot be expected that *NF* proves any strictly local stratified result about familiar mathematical objects which is not also a theorem of ZFC.

Questions of consistency with NF of global choice-like statements such as "the universe is linearly ordered" cannot be resolved by the method used here (at least, not without major changes). One statement which seems to be about big sets can be seen to hold in our models: the power set of any well-orderable set is well-orderable, and more generally, beth numbers are alephs. We indicate the proofs: a relation which one of our models of TTT thinks is a well-ordering actually is a well-ordering, because the models are countably complete; so a well-ordering with a certain support has all elements of its domain sets with the same support (a permutation whose action fixes a well-ordering has action fixing all elements of its domain), and all subsets of and relations on the domain are sets with the same support (adjusted for type differential), and this applies further to the well-ordering of the subsets of the domain which we find in the metatheory. Applying the same result to sets with well-founded extensional relations on them proves the more general result about beth numbers. This form of choice seems to allow us to use choice freely on any structure one is likely to talk about in the usual set theory. It also proves, for example, that the power set of the set of ordinals (a big set!) is well-ordered.

NF with strong axioms such as the Axiom of Counting (introduced by Rosser in [13], an admirable textbook based on *NF*), the Axiom of Cantorian Sets (introduced in [5]) or my axioms of Small Ordinals and Large Ordinals (introduced in my [7], which pretends to be a set theory textbook based on *NFU*) can be obtained by choosing λ large enough to have strong partition properties, more or less exactly as I report in my paper [8] on strong axioms of infinity in NFU: the results in that paper are not all mine, and I owe a good deal to Robert Solovay in that connection (unpublished conversations and [17]).

That NF has α-models for each standard ordinal α should follow by the same methods Jensen used for NFU in his original paper [10]. No model of NF can contain all countable subsets of its domain; all well-typed combinatorial consequences of closure of a model of TST under taking subsets of size $< \kappa$ will hold in our models, but the application of compactness which gets us from TST + Ambiguity to NF forces the existence of externally countable proper classes, a result which has long been known and which also holds in NFU.

We mention some esoteric problems which our approach solves. The Theory of Negative Types of Hao Wang (TST with all integers as types, proposed in [21]) has ω-models; an ω-model of NF gives an ω-model of the theory of negative types immediately. The question of existence of ω models of the theory of negative types was open.

In ordinary set theory, the Specker tree of a cardinal is the tree in which the top is the given cardinal, the children of the top node are the preimages of the top under the map $(\kappa \mapsto 2^{\kappa})$, and the part of the tree below each child is the Specker tree of the child. Forster proved using a result of Sierpinski that the Specker tree of a cardinal must be well-founded (a result which applies in ordinary set theory or in NF(U), with some finesse in the definition of the exponential map in NF(U)). Given Choice, there is a finite bound on the lengths of the branches in any given Specker tree. Of course by the Sierpinski result a Specker tree can be assigned an ordinal rank. The question which was open was whether existence of a Specker tree of infinite rank is consistent. It is known that in NF with the Axiom of Counting the Specker tree of the cardinality of the universe is of infinite rank. Our results in this paper can be used to show that Specker trees of infinite rank are consistent in bounded Zermelo set theory with atoms or without foundation (this takes a little work, using the way that internal type representations unfold in TTT and a natural interpretation of bounded Zermelo set theory in TST; a tangled web as described above would have range part of a Specker tree of infinite rank). A bit more work definitely gets this result in ZFA, and we are reasonably confident that our permutation methods can be adapted to ZFC using forcing in standard ways (in which we are not expert) to show that Specker trees of infinite rank can exist in ZF.

We believe that NF is no stronger than TST + Infinity, which is of the same strength as Zermelo set theory with separation restricted to bounded formulas ([9]). Our work here does not show this, as we need enough Replacement for existence of $\beth_{\omega_1}$ at least. We leave it as an interesting further task, possibly for others, to tighten things up and show the minimal strength that we expect holds.

Another question of a very general and amorphous nature which remains is: what do models of NF (or TTT) look like in general? Are all models of NF in some way like the ones we describe, or are there models of quite a different character? There are very special assumptions which we made by fiat in building our model of TTT which do not seem at all inevitable in general models of this theory.

References

[1] Crabbé, M. [1992a] On NFU. *Notre Dame Journal of Formal Logic* 33, pp 112-119.

[2] Scott Fenton. New Foundations set theory developed in metamath. 2015. `https://us.metamath.org/nfeuni/mmnf.html`

[3] Forster, T.E. [1995] Set Theory with a Universal Set, exploring an untyped Universe Second edition. Oxford Logic Guides, Oxford University Press, Clarendon Press, Oxford.

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

[5] Henson, C.W. [1973a] Type-raising operations in NF. Journal of Symbolic Logic 38 , pp. 59-68.

[6] Holmes, M.R. "The equivalence of NF-style set theories with "tangled" type theories; the construction of omega-models of predicative NF (and more)". *Journal of Symbolic Logic* 60 (1995), pp. 178-189.

[7] Holmes, M. R. [1998] Elementary set theory with a universal set. volume 10 of the Cahiers du Centre de logique, Academia, Louvain-la-Neuve (Belgium), 241 pages, ISBN 2-87209-488-1. See here for an on-line errata slip. By permission of the publishers, a corrected text is published online; an official second edition will appear online eventually.

[8] Holmes, M. R. [2001] Strong Axioms of infinity in NFU. Journal of Symbolic Logic, 66, no. 1, pp. 87-116.
("Errata in 'Strong Axioms of Infinity in NFU' ", JSL, vol. 66, no. 4 (December 2001), p. 1974, reports some errata and provides corrections).

[9] Kemeny, J.G. [1950] Type theory vs. set theory (abstract of Ph.D. thesis). *Journal of Symbolic Logic* 15, p. 78.

[10] Jensen, R.B. "On the consistency of a slight(?) modification of Quine's NF". *Synthese* 19 (1969), pp. 250-263.

[11] Quine, W.V. [1945] On ordered pairs. **Journal of Symbolic Logic** 10, pp. 95-96.

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

[13] Rosser, J. B. [1978] Logic for mathematicians, second edition. Chelsea Publishing.

[14] Russell, Bertrand. *Principles of Mathematics*. Routledge Classics 2010 (originally published 1903).

[15] Russell, B.A.W. and Whitehead, A. N.[1910] *Principia Mathematica.* Cambridge University Press.

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

[17] Solovay, R, "The consistency strength of NFUB", preprint on `arXiv.org`, `arXiv:math/9707207 [math.LO]`

[18] 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 (1953), pp. 972-975.

[19] Specker, E.P. [1962] "Typical ambiguity". *Logic, methodology and philosophy of science*, ed. E. Nagel, Stanford University Press, pp. 116-123.

[20] Tarski, first description of TST

[21] Wang, H. [1952] Negative types.

[22] Wiener, Norbert, paper on Wiener pair

[23] Sky Wilshaw and Yaël Dillies and Peter LeFanu Lumsdaine and others. New Foundations is consistent. GitHub repository. `https://leanprover-community.github.io/con-nf/`.

Initial work was done by a group, as indicated by the authorship; since the beginning of 2023, the work, which now amounts to the majority of what has been accomplished, has been done by Sky Wilshaw alone, and the previous work has been very largely reorganized by her.