Another idea about tangled type theory

Randall Holmes

2/18/2022: added conclusions section and biblio; the former needs some reconciliation (it is from a different version) and the latter requires addition of citations.

Contents

1

1 Development of relevant theories

1.1 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 de 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 $\texttt{type}(`x\textrm{'})$: 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 $\texttt{type}(`x\textrm{'}) = \texttt{type}(`y\textrm{'})$. An atomic membership sentence '$x \in y$' is well-formed iff $\texttt{type}`(x\textrm{'}) + 1 = \texttt{type}(`y\textrm{'})$.

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.

1.1.1 Typical ambiguity

TST exhibits a symmetry which is important in the sequel.

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

If ϕ is a formula, define ϕ^+ as the result of replacing every variable x (free and bound) in ϕ with x^+. 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).

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.

1.1.2 Historical remarks

TST is not the type theory of the *Principia Mathematica* of Russell and Whitehead, 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*. 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 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, 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.

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 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.

1.2 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) 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). 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 $|\iota"X| = |\mathcal{P}(A)|$; 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 $|\iota"X| = |\mathcal{P}(V)|$, because $|\iota"V| < |\mathcal{P}(V)| \leq |V|$, 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 $\iota^{(n-1)-i}"\mathcal{P}^i(X)$. 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 we can define truth for formulas in these natural models of TST_n for each 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.

1.3 New Foundations and NFU

In 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^*, 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 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, refined by others. One obtains a finite axiomatization of NF by analogy with the method of finitely

axiomatizing von Neumann-Gödel-Bernays predicate class theory. It should further be noted that the first thing one does with the 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.

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 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 (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 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

$|\iota\text{``}V| < |(V)| \leq |V|$ (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.

1.4 Tangled type theory TTT and TTTU

In 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 $\text{type}(`x\text{'}) < \lambda$; 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 $\text{type}(`x\text{'}) = \text{type}(`y\text{'})$ and whose atomic sentences '$x \in y$' satisfy $\text{type}(`x\text{'}) < \text{type}(`y\text{'})$, 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 $\mathtt{rng}(s\lceil\{0, \ldots, n-1\}) = A$. 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 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 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.

1.4.1 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.

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 $2^{|\iota^{n+1}\text{“}X|} = \iota^n\text{“}X$, 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.

14

2 A model of TTT

We describe a model of TTT.

2.1 Bare preliminaries: cardinal parameters

We will work in ZFA: it is convenient to have a large collection of atoms.

We specify a limit ordinal λ, whose elements will be used as indices for types. We will build a structure for a type theory in which there are membership relations from each positive type to each higher type (type 0 in this structure is special). The intention is that the positive types of this structure will support an interpretation of TTT.

We specify a regular uncountable cardinal $\kappa > \lambda$: sets of size $< \kappa$ we call small and other sets we call large.

We specify a strong limit cardinal μ which is greater than κ and has cofinality at least κ. There are μ atoms.

2.2 Atoms and litters

The set of atoms is type 0 of the structure.

We provide a partition of type 0 into sets of size κ called litters. We define a near-litter as a set of atoms with small symmetric difference from a litter. For any near-litter, we define $[N]$, the (untyped) local cardinal of N, as the set of all near-litters with small symmetric difference from N. Note that there are μ near-litters. For a near-litter N, we define both N° and $[N]^\circ$ as the unique litter which has small symmetric difference from N and so belongs to $[N]$.

2.3 The elements of the positive types described, with their primary extensions

Elements of a positive type α of the structure are of the form (α, β, B), where B is a subset of type $\beta < \alpha$ (β may be 0).

Any such triple (α, β, B) with $\beta < \alpha$ and B a subset of type β is called a type α class; the actual elements of type α (the type α sets) make up a proper subset of the type α classes: we stipulate at this point only that the cardinality of each type is μ. We may refer to any (α, β, B) which is an element of type α as a *typed set*, when we do not care what the value of α is.

The intention is that the type β elements of the type α class (α, β, B) are the elements of B in the usual sense. We write this more formally thus: we are giving the first clause of the definition of the relation $\in_{\beta,\alpha}$ of interpreted membership of type β sets $(\beta > 0)$ in type α classes $(\alpha > \beta)$, which is, for $0 < \beta < \alpha$ and B a subset of type β, $x \in_{\beta,\alpha} (\alpha, \beta, B)$ iff $x \in B$.

2.4 The other extensions of an element of a positive type: membership relations in the structure defined: typed local cardinals introduced

The other extensions of (α, β, B) are determined using some additional machinery.

For any (near-)litter N, $(\alpha, 0, N)$ will be a type α set (actually belonging to type α) called a typed (near-)litter. We define $[(\alpha, 0, N)]$ as

$$\{(\alpha, 0, M) : M \in [N]\} :$$

this is the typed local cardinal operation, and these objects are the typed local cardinals over type α (subsets of type α, not elements of type α). We define $(\alpha, 0, N)^{\circ}$ and $[(\alpha, 0, N)]^{\circ}$ both as $(\alpha, 0, N^{\circ})$.

The type γ extension of a type α class (α, β, B) where $\gamma \in \alpha \setminus \{0, \beta\}$ will be a union of type γ local cardinals correlated with the elements of B.

To support this, we provide, for each $\beta \geq 0$ and γ positive and distnct from β, an injective map $\chi_{\beta,\gamma}$ from type β into the set of typed local cardinals over type γ, with the stipulation that the ranges of $\chi_{\beta,\delta}$ and $\chi_{\gamma,\delta}$ are disjoint if $\beta \neq \gamma$. These maps can coherently be selected if other aspects of the construction succeed: the extent of the typed local cardinals in each type is evident before the construction starts, and disjoint subsets $\mathtt{rng}(\chi_{\beta,\gamma})$ of cardinality μ of the typed local cardinals over type γ can be reserved for each appropriate β, γ before the construction starts, with each such map actually being selected after type β is defined (and successfully shown to be of size μ)

Now, an element x of type γ belongs to the type γ extension of (α, β, B) in the interpreted sense of our structure just in case it belongs to $\bigcup \chi_{\beta,\gamma} "B$ in the usual sense. That is, when γ is less than α and distinct from 0 and β, we define $x \in_{\gamma,\alpha} (\alpha, \beta, B)$ as holding iff $x \in \bigcup \chi_{\beta,\gamma} "(B)$.

2.5 Extensionality enforced in the type structure

In order to enforce extensionality, certain typed classes have to be excluded from being typed sets. The obstruction to extensionality is that (α, γ, G) and $(\alpha, \beta, \bigcup \chi_{\gamma,\beta} \text{``} G)$ have the same type $\beta$ extension. If $(\alpha, \gamma, G)$ is to be permitted, $(\alpha, \beta, \bigcup \chi_{\gamma,\beta} \text{``} G)$ must be excluded.

The precise description of the classes permitted by the extensionality criterion (not all of these will be typed sets) is that (α, γ, G) is permitted (for $\gamma > 0$) only if G is not for any $\beta < \alpha$ a union of elements of the range of $\chi_{\beta,\gamma}$ or if it is the result of iterated applications of the operation

$$(H \mapsto \bigcup \chi_{\beta,\gamma} \text{``} \bigcup \chi_{\gamma,\beta} \text{``} H)$$

to a G permitted for the first reason. The second clause is needed because otherwise certain extensions would be excluded completely which our eventual exact criterion for sethood will cause us to include.

For this to work properly, we need to ensure that every orbit in

$$(H \mapsto \bigcup \chi_{\beta,\gamma} \text{``} \bigcup \chi_{\gamma,\beta} \text{``} H)$$

can be extended to an orbit which contains an element which is not an image under this operation. It is sufficient for this to (when constructing χ maps $\chi_{\delta,\epsilon}$) use an auxiliary well-ordering of each type of order type μ in which the typed near-litters are placed in advance (because we have to enforce this condition for types not yet constructed) and ensure that $\chi_{\delta,\epsilon}$ maps any typed near-litter N to a local cardinal all of whose elements appear in the order of type ϵ at a later position than N appears in the order of type δ. This prevents any set from having infinitely many iterated preimages under

$$(H \mapsto \bigcup \chi_{\beta,\gamma} \text{``} \bigcup \chi_{\gamma,\beta} \text{``} H)$$

and enables the definition of the classes excluded by the extensionality criterion to be complete.

Note that it is a consequence of this criterion that if $(\alpha, \beta, \emptyset)$ is a set, $\beta = 0$.

This does seem to require a wee bit of explanation. Note that if $(\alpha, \beta, \bigcup \chi_{\gamma,\beta} \text{``} G)$ is to be excluded using a blanket exclusion of sets which are unions of range elements of a χ function (which was our mistaken approach in an earlier draft of this paper) then we would have no way to get a name for its γ extension

$\bigcup \chi_{\beta,\gamma}``\bigcup \chi_{\gamma,\beta}``G$, though we could easily get names for sets close to it (say by omitting one element so that it ceases to be a union of range elements of a χ function), so clearly if we expect our type structure to have sensible properties we cannot exclude this extension, and so we allow it as a γ extension.

We prove that extensionality holds for classes permitted by the extensionality criterion.

Suppose that (α, β, B) and (α, γ, G) are permitted by the extensionality criterion and have the same δ extension. Our goal is to show that $(\alpha, \beta, B) = (\alpha, \gamma, G)$.

If B is empty, then G must be empty, and $\beta = \gamma = 0$, so our goal is achieved.

If $\beta = \gamma$, then whether $\delta = \beta$ or not, $B = G$, so our goal is achieved.

Suppose that B and G are nonempty and $\beta \neq \gamma$.

If δ is not equal to β or γ, then the δ-extension of (α, β, B) is a union of range elements of $\chi_{\beta,\delta}$ and the δ-extension of (α, γ, G) is a union of range elements of $\chi_{\gamma,\delta}$, but this is impossible because these ranges are disjoint and the extensions are nonempty.

So δ must be equal to one of β and γ, without loss of extensionality γ.

So, (α, γ, G) is permitted under the extensionality criterion. The $\delta = \gamma$ extension is G.

Thus the $\delta = \gamma$ extension of (α, β, B) must be $\bigcup \chi_{\beta,\gamma}``B$, and this must be equal to G.

Each of B and G must be obtained by alternating applications of $\bigcup \chi_{\beta,\gamma}``$ and $\bigcup \chi_{\gamma,\beta}``$ as type appropriate from a set which is not the union of range elements of the appropriate χ map, and this is clearly impossible: for each one, satisfying this condition prevents the other from satisfying the condition.

If G is not a union of range elements of $\chi_{\beta,\gamma}$, then $\bigcup \chi_{\beta,\gamma}``B = G$ is impossible.

If G is $\bigcup \chi_{\beta,\gamma}``\bigcup \chi_{\gamma,\beta}``G'$, and G' is not a union of range elements of $\chi_{\beta,\gamma}$, then $B = \chi_{\gamma,\beta}``G'$ is excluded by the extensionality criterion, because $B$ must either fail to be a union of range elements of $\chi_{\gamma,\beta}$ or be a set $\bigcup \chi_{\gamma,\beta}``\bigcup \chi_{\beta,\gamma}``H$, and both are excluded.

And so forth.

2.6 The definition of the positive types

Now we need to declare which type α classes are typed sets. The criterion is a symmetry criterion.

2.6.1 Allowable permutations introduced

We define the notion of α-allowable permutation for $\alpha < \lambda$.

A 0-allowable permutation is a permutation π of the atoms such that for any near-litter N, $\pi``N$ is also a near-litter.

An α-allowable permutation, for positive α, is a permutation of the type α classes with additional properties described in the next paragraph.

We suppose that we have already defined β-allowable permutations for $\beta < \alpha$ with the stipulation that for $0 < \beta < \alpha$, β-allowable permutations send type β sets to type β sets. We then define an α-allowable permutation as a permutation π of type α classes such that for each $\beta < \alpha$ there is a β-allowable permutation π_β such that $\pi(\alpha, \beta, \{b\}) = (\alpha, \beta, \{\pi_\beta(b)\})$ for every b in type β, and for every subset B of type β, $\pi(\alpha, \beta, B) = (\alpha, \beta, \pi_\beta``B)$, with the further condition that type α classes with the same δ-extension for any $\delta < \alpha$ have images under π with the same δ-extension.

The last condition is a coherence condition relating π_β and π_γ for β, γ distinct ordinals less than α. The γ-extension of $(\alpha, \beta, \{b\})$ is $\chi_{\beta,\gamma}(b)$, so this type α class has the same γ-extension (prior to the extensionality adjustment) as the type α class $(\alpha, \gamma, \chi_{\beta,\gamma}(b))$. $\pi(\alpha, \beta, \{b\})$ is $(\alpha, \beta, \{\pi_\beta(b)\})$. $\pi(\alpha, \gamma, \chi_{\beta,\gamma}(b)) = (\alpha, \gamma, \pi_\gamma``\chi_{\beta,\gamma}(b)$, which is to have the same $\gamma$-extension as $(\alpha, \beta, \{\pi_\beta(b)\})$ whence we must have $\pi_\gamma``\chi_{\beta,\gamma}(b) = \chi_{\beta,\gamma}(\pi_\beta(b))$.

This condition can be further analyzed. There is a litter L such that $(\gamma, 0, L) \in \chi_{\beta,\gamma}(b)$ $[(\gamma, 0, L) = \chi_{\beta,\gamma}(b)^\circ$, so $L = \pi_3(\chi_{\beta,\gamma}(b)^\circ)]$. We find that $\pi_\gamma(\gamma, 0, L) = (\gamma, 0, \pi_{\gamma,0}``L) \in \chi_{\beta,\gamma}(\pi_\beta(b))$ so if $(\gamma, 0, M) \in \chi_{\beta,\gamma}(\pi_\beta(b))$ we have $(\pi_{\gamma,0}``L)\Delta M$ small. We can write compactly $\pi_{\gamma,0}``(\pi_3(\chi_{\beta,\gamma}(b)^\circ))^\circ = \pi_3(\chi_{\beta,\gamma}(\pi_\beta(b))^\circ)$. Note that in the last equation I exploit the ability to apply the $-^\circ$ litter selection operation to both typed and untyped near-litters and typed local cardinals: the third projection operator π_3 appears because we need to translate between typed and untyped litters.

[NOTE: I need to verify coherence with the extensionality criterion]

2.6.2 Derived allowable permutations

We define permutations derived from an α-allowable permutation. We define π_A for each finite nonempty subset A of λ whose largest element is α. $\pi_{\{\alpha\}}$ is defined as π; define A_1 as $A \setminus \{\mathtt{min}(A)\}$ if this is nonempty, and define $\pi_A = (\pi_{A_1})_{\mathtt{min}(A)}$ otherwise, noting the hypothesis of the recursion, that π_A is a $\mathtt{min}(A)$-allowable permutation.

2.6.3 Supports and symmetry: type α sets defined as symmetric type α classes

We define an α-support as a small (cardinality $< \kappa$) well-ordering of pairs (A, x) where A is a nonempty finite subset of λ with maximum α, and x is an element of type $\mathtt{min}(A)$ which is an atom if $\mathtt{min}(A) = 0$ and otherwise is a typed near-litter. Where $u = (C, y)$ is in the domain of a support S, we will use the notation $\pi_1(u)$ to refer to A and $\pi_2(u)$ to refer to y: π with these particular suffixes should not be read as a permutation.

The action of an α-allowable permutation π on an α-support S is to replace each (A, x) in its domain with $(A, \pi_A(x))$: the resulting α-support we will denote by $\pi[S]$.

An object X of type α has support S if S is an α-support and any α-allowable permutation π with $\pi[S] = S$ has $\pi(x) = x$.

A type α class is symmetric iff it has an α-support. The elements of type α (the type α sets) are exactly the symmetric type α classes which are permitted by the extensionality criterion.

Note that the image of an α-symmetric set under an α-allowable permutation is readily seen to be α-symmetric, and typed near-litters of type α are obviously α-symmetric, covering another previous assertion of typed sethood. NOTE: we must demonstrate that classes permitted by the extensionality criterion are mapped to classes permitted by the extensionality criterion.

2.7 Predicative TTT comprehension in the type structure

The coherence condition ensures that for any α-allowable permutation π, $x \in_{\beta,\alpha} y$ iff $\pi_\beta(x) \in_{\beta,\alpha} \pi(y)$. For each $\gamma \in A$, define $A_{(\gamma)}$ as the upper segment of A whose smallest element is γ. If ϕ is a formula in which every

type which occurs is in a fixed finite subset A of λ and each membership formula $x \in_{\gamma,\delta} y$ which occurs has $\gamma < \delta$ in A with no intervening elements of A, application of the result of the previous sentence allows us to see that the truth value of ϕ is unchanged if each membership statement $x \in_{\gamma,\delta} y$ is replaced with $\pi_{A_{(\gamma)}}(x) \in \pi_{A_{(\delta)}}(y)$: of course, systematically replacing each x of type γ with $\pi_{A_{(\gamma)}}(x)$ also leaves equations unaffected in truth value. It follows from this that if x is a type β variable (second highest type in A) and y is a type α variable (highest type in A) then $(\alpha, \beta, \{x : \phi\})$ (where ϕ satisfies the conditions described above) is a symmetric type α class. The reason is that any permutation π such that for each γ, $\pi_{A_{(\gamma)}}$ fixes each type γ parameter in ϕ will fix this set, and this gives us a description of an α-support for it. Thus we have the type system satisfying the interpretation of a predicative version of the comprehension axiom of tangled type theory. Note that it is sufficient to get the full comprehension axiom of tangled type theory to show that the axiom of union (in each typed version) is satisfied.

2.8 Requirements still to be met

An obvious requirement is that we need to show that each type is of cardinality μ, or the selection of the χ maps will fail.

As just noted, we need to show that the axiom of set union holds in each typed version in the structure.

We address both of these points by careful analysis of the symmetry of the structure.

2.9 Strong supports introduced

We define an A-support as a support in which each domain element (B, x) has B a not necessarily proper downward extension of A. With each A-support S, we associate the $\mathtt{min}(A)$-support S/A:

$$(B, x)\, S\, (C, y) \leftrightarrow (B \setminus A_1, x)\, S/A\, (C \setminus A_1, y)$$

We define a strong support S as a α-support in which

1. each element (B, x) of the domain has x an atom or a litter [any support of an object X can be transformed into a support satisfying this condition by replacing each domain element (B, x) with $x = (\beta, 0, N)$ a typed near-litter and not a typed litter with $(B, (\beta, 0, N^\circ))$ together with all $(B \cup \{0\}, z)$ with $z \in N \Delta N^\circ$],

21

2. each domain element (B, x) with x an atom and $B \neq \{0\}$ (equivalently, $\alpha \neq 0$) is preceded by an element (B_1, L) where $\pi_3(L)$ is the litter containing x,

3. and each element (B, x) where $[x] \in \mathtt{rng}(\chi_{\delta, \mathtt{min}(B)})$ for some $\delta < \mathtt{min}(B_1)$ has the segment it determines in S including a $B_1 \cup \{\delta\}$ support T such that $T/(B_1 \cup \{\delta\})$ is a δ-support of $\chi_{\delta, \mathtt{min}(B)}^{-1}([x])$.

   Notice that this further implies that each element (B, x) where $[x] \in \mathtt{rng}(\chi_{\delta, \mathtt{min}(B)})$ for some $\delta < \mathtt{min}(B_1)$ [note that this does require $\mathtt{min}(B) < \alpha$] has the segment it determines in S including a strong $B_1 \cup \{\delta\}$ support T^* of $\chi_{\delta, \mathtt{min}(B)}^{-1}([x])$, by which we mean a support T^* such that $T^*/(B_1 \cup \{\delta\})$ is a strong δ support of $\chi_{\delta, \mathtt{min}(B)}^{-1}([x])$.

Every object of type α has a strong α-support. The strategy for constructing one is to modify the support through ω stages, at each stage inserting required domain elements with second projections litters (immediately before elements of the domain with atomic second projections which cause them to be required) and required supports (immediately before domain elements with litter second projections which cause them to be required, assuming that near-litters which are not litters are already handled as described above), duplicate items being handled by removing all but the first occurrence. This process will terminate with a strong α-support because adjustments can be made only finitely many times at any given point: an adjustment adding a litter will be made only once, and a $B_1 \cup \{\delta\}$-support added to deal with an element (B, x), x a litter will have $\delta < \mathtt{min}(B_1) < \alpha$, and any support required to be added to deal with elements of this δ-support will be an E-support for some E with $\mathtt{min}(E) < \delta$: there cannot be an indefinite regress of new supports required to be added.

2.10 Local approximations to allowable permutations and the freedom of action theorem

A α-local approximation is a function π^0 such that if $\pi^0(A, x)$ is defined, A is a nonempty finite subset of λ with maximum α and minimum 0 and x is an atom, and the set of second projections of elements of the range of the domain of π^0 has small intersection with each litter, and each map $\pi_A^0 = (x \mapsto \pi^0(A, x))$ has an inverse and has domain the same as its range.

   We prove a

Theorem (Freedom of Action): For any α-local approximation π^0 there is an allowable permutation π such that $\pi_A(x) = \pi^0(A, x)$ whenever the latter is defined, and further, when x belongs to the litter L and either $\pi_A(x) \notin (\pi_A \text{``} L)^\circ$ or $\pi_A^{-1}(x) \notin (\pi_A^{-1} \text{``} L)^\circ$ (such an x is called an exception of π_A), we also have that (A, x) is in the domain of π^0.

Proof: The freedom of action theorem is proved for a given α-local approximation π^0 by constructing the desired π by recursion along strong supports.

We are given as data before the process starts a well-ordering $<_L$ of each litter L of order type κ.

We are given as data before the process starts, for each A with maximum α and $\mathtt{min}(A) > 0$, a permutation ξ_A of all typed litters x over type $\mathtt{min}(A)$ for which $[x]$ is not in the range of any $\chi_{\delta,\mathtt{min}(A)}$ for $\delta < \mathtt{min}(A_1)$.

Suppose that in a strong support S we have already computed $\pi_C(z)$ for each (z, C) before a given (x, A) in the order S.

If x is an atom and $\pi^0(A, x)$ is defined, set $\pi_A(x)$ to $\pi^0(A, x)$.

Otherwise, if x is an atom we have by the hypothesis of the recursion already computed $\pi_{A_1}(L)$, where L is the litter containing x. Let $\pi_A \lceil L$ be the union of the restriction of π_A^0 to L and the unique map f from $L \setminus \mathtt{dom}(\pi_A^0)$ onto $\pi_{A_1}(L)^\circ \setminus \mathtt{dom}(\pi_A^0)$ such that if $u <_L v$, $f(u) <_{\pi_{A_1}(L)^\circ} f(v)$, and set $\pi_A(x) = \pi_A \lceil L(x)$.

There remains the case where x is a typed litter. If $[x]$ is not in the range of any $\chi_{\delta,\mathtt{min}(A)}$ for $\delta < \mathtt{min}(A_1)$, then set

$$\pi_3(\pi_A(x)) = \pi_{A\cup\{0\}}^0 \text{``} \pi_3(x) \cup (\pi_3(\xi_A(x)) \setminus \mathtt{rng}(\pi_{A\cup\{0\}}^0)).$$

Of course the first two projections of $\pi_A(x)$ are the same as the corresponding projections of x.

If $[x]$ is in the range of $\chi_{\delta,\mathtt{min}(A)}$ for $\delta < \mathtt{min}(A_1)$, more exciting things happen. In this case we know that there is a support T included in S before (x, A) such that $T/(A_1 \cup \{\delta\})$ is a δ-support (which we can suppose a δ-strong support because S is strong) of $\chi_{\delta,\mathtt{min}(A)}^{-1}([x])$. For every (z, C) in the domain of $T/(A_1 \cup \{\delta\})$, we have already defined $\pi_{A_1 \cup C}(z)$ by inductive hypothesis. Extend the map sending each such (C, z) with z an atom to $\pi_{A_1 \cup C}(z)$ to a δ-local approximation π_*^0, in

23

such a way that for any typed litter v with (D, v) in the subsupport, for which we have already computed $\pi_{A_1 \cup D}(v)$, and any $(D \cup \{0\}, u)$ in the domain of the δ-local approximation with u an atom, the action assigned to the δ-local approximation at $(D \cup \{0\}, u)$ is consistent with the known action of $\pi_{A_1 \cup D}$ on (D, v), and with the additional stipulation that the maps ξ_C that we use in the construction of π_* coincide with the maps $\xi_{A_1 \cup C}$ provided for the construction of π. The need to extend the local approximation arises from the need to fill out orbits of maps $\pi^0_{*,C} = (v \mapsto \pi^0_*(C, v))$ so that they are injective and have domain the same as their range: no more than countably many new domain values per initially given domain value are needed, and it should be evident that all requirements can be met.

Extend this map to a δ-allowable permutation with no exceptions other than domain elements (we can do this by an inductive hypothesis because $\delta < \alpha$), which we call π_*. $(\pi_*)_C(z)$ for z an atom with (C, z) in the δ-support is $\pi^0_*(C, z) = \pi_{A_1 \cup C}(z)$; we claim that $(\pi_*)_C(z) = \pi_{A_1 \cup C}(z)$ is also true for z a typed litter appearing as second projection of an element (C, z) of the δ-support. If there is a (C, z) in $T/(A_1 \cup \{\delta\})$ such that $(\pi_*)_C(z) \neq \pi_{A_1 \cup C}(z)$ there is a first one, which we briefly call (C, z).

If $[z]$ is not in the domain of $\chi_{\epsilon, \min(C)}$ for any $\epsilon < \min(C_1)$, then $(\pi_*)_C(z)$ belongs to the same local cardinal $[\xi_C(z)])$ as the value we have computed for $\pi_{A_1 \cup C}(z)$ (by the stipulation on ξ maps used in the construction of π_* stated above). In fact, these two values must be the same. $\pi_{A_1 \cup C \cup \{0\}}$ maps elements of $\pi_3(z)$ out of $\pi_3(\pi_{A_1 \cup C}(z))^\circ$ or non-elements of $\pi_3(z)$ into $\pi_3(\pi_{A_1 \cup C}(z))^\circ$ only if this is directly dictated by the local approximation π^0. New elements of the domain and range of π^0_* will not enforce an action on an atom in $\pi_3(z)$ in which conflicts with the action of π_C on the typed litter z, by construction. So the action on the litter z of the two maps must be the same: any difference would imply the presence of an exception of a derivative of π_* or π not derived from the domain of the corresponding local approximation.

If $[z]$ is in the domain of $\chi_{\epsilon, \min(C)}$ for a $\epsilon < \min(C_1)$, then the actions of appropriate derivatives of π and π_* agree on a support of $\chi^{-1}_{\epsilon, \min(C)}([z])$, and so (recalling that π_C has already been computed using an actual allowable permutation extension at z, just as we are doing with π_A at x) the values returned by the two functions at z can only have small

symmetric difference (the local cardinals of the values must be the same). Exactly as in the previous paragraph, $\pi_{C\cup\{0\}}$ maps elements of $\pi_3(z)$ out of $\pi_3(\pi_C(z))^\circ$ or non-elements of $\pi_3(z)$ into $\pi_3(\pi_C(z))^\circ$ only if this is directly dictated by π^0. New elements of the domain and range of π_*^0 will not enforce an action on an atom in $\pi_3(z)$ which conflicts with the action of π_C on the typed litter z, by construction. So the action on the litter z of the two maps must be the same: any difference would imply the presence of an exception of a derivative of π or π_* not derived from the domain of the corresponding local approximation.

It follows that π_* acts as $\pi_{A_1\cup\{\delta\}}$ does on a support of $\chi_{\delta,\mathtt{min}(A)}^{-1}([x])$, which determines how π_A must act on x up to small symmetric difference, so we can define $\pi_3(\pi_A(x))$ as

$$\pi_{A\cup\{0\}}^0 \text{``}\pi_3(x) \cup (\pi_3(\chi_{\delta,\mathtt{min}(A)}(\pi_*(\chi_{\delta,\mathtt{min}(A)}^{-1}([x])))^\circ) \setminus \mathtt{rng}(\pi_{A\cup\{0\}}^0)).$$

Finally, we need to argue that this process will give the same result for any value $\pi_A(x)$ independently of the strong support S used for the procedure. If this fails, then in some support S there will be a first domain element x at which a different value could be obtained by computation along a different support T. Merge the segment S_x (the segment of S before x) and T_x suitably to discover that the same value must be computed along both S_x and T_x at the purported first bad element. The suitable merger is to put T_x before S_x and eliminate any duplicates by retaining the first occurrence. Computations along T_x will of course be the same and force the same value at x, and this means that the values in S_x appearing before x are compatible with computing this same value at x, which is a contradiction.

The fact that no derivative of π has any exceptions not determined by values of π^0 is transparently enforced by the recursive procedure.

2.11 Analysis of orbits to determine cardinalities of types

Each type is clearly of cardinality at least μ, since it contains all typed near-litters of that type, and there are μ of these.

We need to show that each positive type is of cardinality exactly μ, in order to ensure that we can actually select the χ maps at each stage.

Type 0 is of cardinality μ, of course. We fix a positive α and assume that we have already established that each type with index $\beta < \alpha$ is of cardinality μ, which allows us to define type α (all required χ maps required to define α-allowable permutations thus being possible to select).

We define a nice β-support as a support $\pi[S]$ where S is a strong β-support and π is a β-allowable permutation. The principal difference is that domain elements (B, x) of a nice support can have x a typed near-litter which is not a litter.

For any object x of type β with a strong support S, we define a function $\zeta_{x,S}$ which we call a β-coding function for x such that $\zeta_{x,S}(\pi[S]) = \pi(x)$ for each β-allowable permutation. Notice that if π' is a β-allowable permutation and we have $\pi[S] = \pi'[S]$ we have $\pi(x) = \pi'(x)$ because S is a support.

The domain of a coding function is an orbit under β-allowable permutations in the action of such permutations on β-supports, including at least one strong support.

Our ultimate claim is that for each $\beta \leq \alpha$ there are $< \mu$ β-coding functions, and so there are no more than μ type α sets, since each one can be determined by an α-coding function and an α-support, and there are μ α-supports. We can assume that there are $< \mu$ β-coding functions for $\beta < \alpha$ as part of our inductive hypothesis.

It is clear at the basis that there are $< \mu$ 0-coding functions: a 0 support is simply a small well-ordering of pairs $(\{0\}, x)$ for atoms x and the coding functions are determined by ordinals $< \kappa$: each coding function is determined by the order type of its domain elements $(< \kappa)$ and an ordinal γ less than this order type and returns x, where $(\{0\}, x)$ is the γ'th element of the support.

With each nice support, associate a combinatorial object called an *orbit specification* (we will show that it actually determines the orbit in the allowable permutations in which the support lies).

The orbit specification of a support S is a well-ordering $o(S)$ of the same length as S. Corresponding to an item (B, x) in the domain of S at ordinal position γ [for which we use the notation S_γ], we describe the item at position γ in the orbit specification $o(S)$.

1. If x is an atom, this item is $(\gamma, 0, B, \delta)$ where $\delta < \gamma$ and $\pi_2(S_\delta)$ is a near-litter to which x belongs.

2. If x is a near-litter with $[x]$ not belonging to any $\mathtt{rng}(\chi_{\delta,\mathtt{min}(B)})$ with $\delta < \mathtt{min}(B_1)$, this item is $(\gamma, 1, B)$.

3. If x is a near-litter with $[x]$ belonging to the range of $\mathtt{rng}(\chi_{\delta,\mathtt{min}(B)})$ with $\delta < \mathtt{min}(B_1)$, let T be the maximal $B_1 \cup \{\delta\}$-support of $\chi_{\delta,\mathtt{min}(B)}^{-1}([x])$ included in the initial segment of S of length γ [T will be nice and can be determined just by looking at first projections of domain elements of S, information which is contained in $o(S)$], and let $o(S)_\gamma = (\gamma, 2, B, \zeta_{\chi_{\delta,\mathtt{min}(B)}^{-1}([x]),T/(B_1 \cup \{\delta\}})$

We argue that the orbit specification of S determines the orbit in the allowable permutations of appropriate index in which S lies. One side of this is direct: it is straightforward to verify that the action of an allowable permutation π on S will send S to a support $\pi[S]$ with the same orbit specification.

It is trickier to establish that if S and T are supports with the same orbit specification, that there is an allowable permutation π such that $\pi[S] = T$. The strategy is to construct a local approximation to such a π and use the Freedom of Action theorem to establish its existence.

Where $\pi_2(S_\gamma)$ is an atom, we set $\pi_0(S_\gamma) = \pi_2(T_\gamma)$ (note that S_γ is a pair of the right kind to feed to this local approximation!)

Where $\pi_2(S_\gamma)$ is a near-litter not in the domain of an appropriate χ map, we arrange for the ξ map of the appropriate index to map $\pi_2(S_\gamma)^\circ$ to $\pi_2(T_\gamma)^\circ$.

Where $\pi_2(S_\gamma)$ is a typed near-litter $(\beta, 0, N)$ then $\pi_2(T_\gamma)$ will be a typed near-litter $(\beta, 0, N')$. This drives addition of pairs $(\pi_1(S_\gamma), x)$ to the domain of π^0 to be mapped to values y, with the effect that each element x of $N \setminus N^\circ$ is mapped into N', and each element y of $N' \setminus N'^\circ$ is the target of an x in N, then fill in countable orbits in appropriate maps π_B^0 for each new value in ways which will not create exceptions as we aim to map each $\pi_2(S_\gamma)$ to $\pi_2(T_\gamma)$.

Then apply the Freedom of Action theorem to get the desired π. It maps atoms which are second projections of domain elements of S to corresponding atoms which are second projections of domain elements of T in the same position by construction. It maps near-litters which are second projections of domain elements of S to corresponding near-litters which are second projections of domain elements of T in the same position: consider the first second projection of a domain element of S which is an exception. Either consideration of a ξ map or consideration of the action already determined on a support of the local cardinal of the near-litter by the inverse of an appropriate χ map ensures that the near-litter second projection of a domain of S is mapped to a near-litter with the same local cardinal as the desired target.

But then the fact that π has no exceptions not determined by elements of π^0 ensures that it is in fact mapped to exactly the correct near-litter which is second projection of the corresponding domain element of T.

Notice that there are no more than μ orbit specifications, so there are no more than μ orbits in the action of allowable permutations on nice supports.

Now we give a formal specification of α-coding functions which makes it clear how many of these there are. Let x be a type α set (α, β, X) with strong α-support S. For each $y \in X$, choose a strong β-support T_y which includes all the (C, z)'s for which $(C \cup \{\alpha\}, z) \in S$ as an initial segment in the order inherited from S. We claim that the support S and the set of all ζ_{y,T_y} determine x and so determine the coding function $\zeta_{x,S}$. The key claim is that X is exactly the set of all $\zeta_{y,T_y}(U)$ where $o(U) = o(T_y)$ and U has the same initial segment taken from S that T_y does. Any element y of X is $\zeta_{y,T_y}(T_y)$ and T_y trivially meets the conditions on U. For a general $\zeta_{y,T_y}(U)$, we construct supports T_y^* and U^* by adding α to each first projection in both supports and expanding both supports to have S as an initial segment. As above, build an α-allowable permutation π whose action sends T_y^* to U^*. π fixes x because its action fixes S, and π_β maps y to $\zeta_{y,T}(U)$, so $\zeta_{y,T}(U) \in X$.

So each type α set is determined by a strong support and a set of β-coding functions for $\beta < \alpha$. There are $< \mu$ such sets of coding functions because μ is strong limit. So there are μ type α sets.

Now we have to argue that there are actually $< \mu$ coding functions for type α sets (to show that type α is the right size, it is sufficient to see that there are μ of them, but the stronger result is needed for the argument to keep going). There are $< \mu$ α-orbit specifications, these being small well-orderings of objects taken from sets known to be of size $< \mu$ (by inductive hypothesis in the case of coding functions). We show that we can define the function $\zeta_{x,S}$ using just $o(S)$ and the set of ζ_{y,T_y}'s for $y \in X$: $\pi_3(\zeta_{x,S}(S'))$ is the set of all $\zeta_{y,T_y}(U)$ where U has the initial segment taken from S' analogous to the initial segment taken from S which appears in T_y, for any S' in the orbit of S under allowable permutations. And this makes it clear that there are $< \mu$ α-coding functions.

And this completes the argument by induction that all types are of size exactly μ, which completes the verification that all types can actually be constructed.

2.12 The verification of the axiom of union

We now want to show that the axiom of union holds in all of its various typed versions in the type structure we have defined. It is sufficient to show that unions of sets of singletons are sets along any path through the types (this is simply a fact about predicative type theory). The precise result needed is that for any types $\alpha > \beta > \gamma$, if (α, β, B) is a type α set where each element of B is of the form $(\beta, \gamma, \{x\})$ for some x in type γ, then $(\beta, \gamma, \bigcup\{\pi_3(b) : b \in B\})$ is a type β set.

Because (α, β, B) is a type α set, there is a support S such that for any α-allowable permutation π such that $\pi[S] = S$ we have $\pi(\alpha, \beta, B) = (\alpha, \beta, \pi_\beta "B) = (\alpha, \beta, B)$

We want to find a support T such that for any β-allowable permutation π with $\pi[T] = T$, we have $\pi(\beta, \gamma, \bigcup\{\pi_3(b) : b \in B\}) = (\beta, \gamma, \bigcup\{\pi_3(b) : b \in B\})$.

Let T be obtained from S by $T = \{(A, x) : \mathtt{max}(A) = \beta \wedge (A \cup \{\alpha\}, x) \in S\}$.

Let $\pi[T] = T$. We want to show that $\pi(\beta, \gamma, \bigcup\{\pi_3(b) : b \in B\}) = (\beta, \gamma, \bigcup\{\pi_3(b) : b \in B\})$.

Let $(\beta, \gamma, \{x\})$ be an element of B. Let U be a strong support of x such that each (A, u) with $(A \cup \{\beta\}, u) \in T)$ belongs to the domain of U and these domain elements make up an initial segment of U with

$$(A, x)\, U\, (A', v) \wedge (A' \cup \{\beta\}, v) \in \mathtt{dom}(T) \to (A \cup \{\beta\}, x)\, T\, (A' \cup \{\beta\}, v).$$

This can be arranged by taking any strong support of x and prepending T modified by dropping β from first projections and restricting to elements with γ as largest element of their first projection.

Now use the Freedom of Action theorem to find an α-allowable permutation π' such that $\pi'[S] = S$ and the action of π on domain elements of T correlates with the action of π'_β on corresponding elements of U (so $\pi_A(u) = \pi'_{A \cup \{\alpha\}}(u)$ if $(A \setminus \{\beta\}, u) \in \mathtt{dom}(U)$ and $(A \cup \{\alpha\}, u) \in \mathtt{dom}(T)$. The two requirements are compatible because π fixes u if $(A, u) \in T$ and $(A \cup \{\alpha\}, u) \in S$. These stipulations are made originally for elements of the domains of supports which have second projection atoms, but they follow by inductive arguments in a style already exhibited for elements whose second projections are near-litters.

Our conditions are sufficient to show that π' fixes (α, β, B), so $\pi'_\beta "B = B$, so $\pi'_\beta(\beta, \gamma, \{x\}) = (\beta, \gamma, \{\pi'_{\beta,\gamma}(x)\})$ and we further have this equal to $(\beta, \gamma, \{\pi_\gamma(x)\})$ because of the correlation of actions on T and U and so equal to $\pi(\beta, \gamma, \{x\})$, and we see that this belongs to B. Similarly $(\pi'_\beta)^{-1}(\beta, \gamma, \{x\}) =$

$\pi^{-1}(\beta, \gamma, \{x\})$, and this belongs to B. And this shows that $\pi(\beta, \gamma, \bigcup\{\pi_3(b) : b \in B\}) = (\beta, \gamma, \bigcup\{\pi_3(b) : b \in B\})$.

And this completes the argument.

3 Conclusions, extended results, and questions

[I have copied in the conclusions section of an older version, but what it says should be about right, and may require some revisions to fit in this paper. I also added the bibliography, which again is probably approximately the right one.]

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

NF has no locally interesting combinatorial consequences. Any fact about sets of a bounded standard size which holds in ZFCA 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 ZFCA 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 result about familiar mathematical objects which is not also a theorem of ZFCA (or even 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).

NF with strong axioms such as the Axiom of Counting (introduced by Rosser in [9], an admirable textbook based on *NF*), the Axiom of Cantorian Sets (introduced in [2]) or my axioms of Small Ordinals and Large Ordinals (introduced in my [4] 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 [5] on strong axioms of infinity in NFU: the results in that paper are not all mine, and I owe a good deal to Solovay (unpublished conversations and [11]).

That NF has α-models for each standard ordinal α should follow by the same methods Jensen used for NFU in his original paper [7]. 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 [14]) has ω-models; an ω-model of NF gives an ω-model of TST immediately. This question 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 show that Specker trees of infinite rank are consistent in ZFA [this requires discussion in the tangled type theory approach used here, but is still true]. We are confident that our permutation methods can be adapted to ZFC using forcing in standard ways 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. Our work here does not show this, as we need enough Replacement for existence of $\beth_{\omega_2}$ at least. We leave it to 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 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?

References

[1] 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.

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

[3] 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.

[4] 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.

[5] 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).

[6] Jech, Thomas, *Set theory*, Academid Press 1978, pp. 199-201.

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

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

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

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

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

[12] 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.

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

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