Prefatory material to a proof of Con(NF)

M. Randall Holmes

April 25, 2022

Abstract

We will present preliminary definitions and results which should suggest the nature of the author's claimed and not very approachable proof of the consistency of Quine's "New Foundations". This will include basic definitions, and indication of the fundamental results obtained by Specker, Jensen's proof of the consistency of the related theory *NFU*, and the results of the author (1995) on equiconsistency of NF with "tangled type theory", followed by a mere sketch of the strategy of a version of the proof the author has claimed since 2010, if time permits.

A very basic theory

We define TSF (typed stratified foundations) as the first order theory with sorts indexed by the natural numbers and primitive binary relations of equality and membership, with well-formedness conditions on atomic formulas briefly indicated in the formats $x^i = y^i$ and $x^i \in y^{i+1}$, and the axiom scheme introduced shortly.

We do not always affix numerical superscripts to variables to indicate type. One should be aware that any variable x in a formula does have a type whose index we will write $\texttt{type}(x)$.

The one and only axiom scheme of TSF consists of the universal closures of formulas of the shape

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

where A is not free in ϕ.

Of course TSF is an anachronism (its untyped stratified analogue SF was proposed by Marcel Crabbé in the context of a proof of cut-elimination). One is definitely expecting an identity criterion for the objects of the theory here. They seem to be sets (at least some of them), and we would expect sets with the same elements to be equal.

It is useful to see that TSF can interpret the more natural theory TSTU in which we postulate extensionality for sets.

Define $x =_{\text{TSTU}} y$ as $(\forall z : z \in x \leftrightarrow z \in y)$ if the common type of x and y is positive, and otherwise simply as $x = y$. This is the relationship of coextensionality.

Define $\text{set}(x)$ as $(\forall y \in x : (\forall z =_{\text{TSTU}} y : z \in x))$, for x of positive type. An object x is a set iff

its extension is a union of equivalence classes under coextensionality.

Define $x \in_{\mathrm{TSTU}} y$ as $x \in y \wedge \mathtt{set}(y)$, where the type of y is the successor of the type of x (a condition forced by the form of the definition).

We have constructed an interpretation of a theory TSTU (with equality, membership and sethood relations primitive) which has. in addition to the comprehension axiom of SF already given, the following axioms:

sethood: $x \in y \rightarrow \mathtt{set}(y)$

(weak) extensionality:

$$\mathtt{set}(x) \wedge \mathtt{set}(y) \rightarrow (x = y \leftrightarrow (\forall z : z \in x \leftrightarrow z \in y))$$

set notation: There is exactly one witness to any formula

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

(A not free in ϕ) which is in addition a set, and this we can denote by $\{x : \phi\}$.

Such monsters as the Russell class $\{x : x \notin x\}$ are averted by the fact that $x \notin x$ simply isn't a well-formed formula of our language.

The theory TST (the simple typed theory of sets) is obtained by strengthening the axiom of extensionality:

extensionality:

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

The anachronism here is of course that TST is the original version of this theory. It is certainly not the type theory of Russell: it seems to have been intimated by Norbert Weiner in 1914 and first formally described by Tarski in 1930. In fact TSTU is prior to TSF. Crabbé proposed SF as a system to effect the relative consistency proof for weak extensionality which

I describe (in the untyped context), and also has proved cut elimination for SF as a higher order logic.

We have our reasons for keeping the intellectual priority of these systems inverse to the historical priority.

Typical ambiguity

This theory exhibits a significant symmetry (Russell noticed a similar symmetry in his type theory but it is purer here).

Postulate a injective map $(x \rightarrow x^+)$ which maps the variables of each type i onto the variables of type $i + 1$.

We can then define ϕ^+ as the result of replacing each variable x in ϕ with x^+.

Notice that for any formula ϕ, ϕ^+ is also a formula. For any axiom ϕ, ϕ^+ is also an axiom. And in fact for any theorem ϕ, ϕ^+ is also a theorem. This is true for any of the theories discussed. Any construction we define using notation $\{x : \phi\}$ has an exact analogue in the next higher type of the form $\{x : \phi^+\}$, and so analogues in every higher type.

It is not necessarily the case that ϕ is a theorem if ϕ^+ is a theorem: we have more information about higher types than we do about lower types.

But it is possible to conjecture that $\phi \leftrightarrow \phi^+$ might be true for each sentence ϕ (this is called the axiom scheme of ambiguity) and this is a way of approaching the motivation of New Foundations.

New Foundations

In 1937, Quine proposed New Foundations (NF), a first order unsorted theory with equality and membership as primitives and with axioms of strong extensionality (as above, but with distinctions of type suppressed) and each comprehension axiom of TST with distinctions of type suppressed.

Notice that this does not give $\{x : x \notin x\}$, because this is not a type-erased version of a comprehension axiom of TST, but it does for example give $\{x : x = x\}$, the universal set.

Quine pushed past the hypothesis that the scheme of ambiguity that I describe above might hold to the apparently stronger hypothesis that the types are simply the same and the distinction between them can be suppressed.

Stratification

It is odd to describe the comprehension axiom of one theory in terms of the language of another. It is traditional to say that NF provides for the existence of sets $\{x : \phi\}$ when there is a function σ from variables (as bits of text) to natural numbers 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 $\sigma(x) + 1 = \sigma(y)$. Such a function σ is called a stratification and a formula which admits a stratification is said to be stratified.

We invite the reader to convince themselves that this is the same as my description of the comprehension axiom above. Also, it should be noted that the stratified comprehension scheme is equivalent to the conjunction of finitely many of its instances (obtained in a rather obfuscated form by Hailperin, and later by other workers in nicer forms, by analogy with the finite axiomatization of the von Neumann-Gödel-Bernays predicative class comprehension axiom). The details of this are unimportant, but it is important to notice that it can be done, because such a finite axiomatization does not need to make any mention of types at all.

In any case, Quine's intuition about the motivation of NF was justified by work of Specker. In 1962, Specker showed that NF is equiconsistent with TST $+$ the ambiguity scheme, and also with existence of a model of TST with an isomorphism to its submodel consisting of the types of positive index (the submodel having type indices reduced by 1). The models of these theories will satisfy the same sentences in the natural sense (mod addition or removal of type indices).

These results, though originally stated for TST and NF, adapt to TSTU and NFU, and to TSF and its untyped analogue SF.

Specker had already proved, in 1954, that NF disproves the Axiom of Choice. This means that NF proves Infinity, as the negation of Infinity certainly implies Choice. This cast doubt on the consistency of NF and its motivation:

after all, TST $+$ AC has the same typical ambiguity properties as TST, but by Specker's two results taken together, it must disprove the ambiguity scheme.

Tangled type theory

I present a somewhat different (perhaps alarmingly different) type theory. This theory is due to the author in 1995. I am going to use it to develop a version of Jensen's proof of Con(NFU); the further advantage is that this method of proving Con(NFU) suggests a strategy for attacking Con(NF) [a strange one] as Jensen's original format does not.

TTTU (tangled type theory with urelements) is a first order theory with sorts indexed by the natural numbers, with equality and membership as primitive notions, and with atomic sentences in the formats $x^i = y^i$ and $x^i \in y^j$ $(i < j)$. We handle sethood by providing predicates set^j which identify sets of type j objects in each higher type.

Provide an injective map $(x \mapsto x^+)$ taking the variables of each type i onto the variables of

type $i + 1$, and define x^{+^j} in the obvious way so that $(x \mapsto x^{+^j})$ is injective from each type i onto type $i + j$.

For any formula ϕ and strictly increasing sequence s of natural numbers, and formula ϕ of the language of TSTU, define ϕ^s as the result of replacing each variable x^i in ϕ with $x^{+^{s(i)-i}}$ (and each occurrence of $\mathtt{set}^i$ with $\mathtt{set}^{s(i)}$) (and labelling sethood predicates with the superscript appropriate to the type of their arguments).

The axioms of TTTU are exactly the formulas ϕ^s where s is a strictly increasing sequence of natural numbers and ϕ is an axiom of TSTU.

One thing to realize is that this is not cumulative type theory. The types are disjoint, or at least we do not say anything about how they

might be related to each other ($x^i = y^j$ is ill formed for $i \neq j$, not false (or true)).

The extensionality axioms say, for example that the identity of a type $j > i$ object which is a set (of type i objects) is determined by the collection of its type i elements for each distinct value of i (a type 4 object is completely determined by its type 3 elements if it has any, and also completely determined by its type 2 elements, if it has any).

Equivalence of TTTU and NFU, in some sense

We show that consistency of TTTU and consistency of NFU are equivalent, where NFU is exactly as NF but with weak extensionality.

Suppose there is a model of NFU with domain D and membership relation E. Then there is a model of TTTU in which type i is implemented as $D \times \{i\}$, the membership relation $E_{i,j}$ for $i < j$ is $\{((x,i),(y,j)) : x \, E \, y)$ and $\mathtt{set}^i$ is $\{(x,i) : \mathtt{set}(x)\}$.

Now suppose there is a model of TTTU. We give an argument closely based on Jensen's original proof of Con(NFU) for the consistency of TSTU + Amb, from which the consistency of NFU follows by Specker's result of 1962.

We choose a finite set Σ of formulas of the language of TST. Let $n - 1$ be the largest

type mentioned in Σ. We define a partition of $[\mathbb{N}]^n$ as follows: a set A of n natural numbers goes into a compartment determined by the truth values of the sentences ϕ^s in the model of TTTU, for each $\phi \in \Sigma$ and any fixed s such that the range of the restriction of s to $\{0, \ldots, n-1\}$ is A. The partition has at most $2^{|\Sigma|}$ compartments. By Ramsey's theorem it has a homogeneous set H. Let h be any increasing sequence whose range is contained in H. The set of sentences ϕ^h for ϕ in the language of TST which are true in the model of TTTU give a complete theory of a model of TSTU with the ambiguity scheme restricted to formulas in Σ, so this theory is consistent. By compactness, TSTU + Amb is consistent, and so is NFU by Specker's results.

We can get stronger results by extending the types of TTTU to elements of a larger limit ordinal λ which may have stronger partition properties.

NFU is consistent

That NFU is consistent follows at once, because it is not difficult to present a model of TTTU, however weird this may seem.

Let λ be a limit ordinal.

Let type $\alpha < \lambda$ be represented by $V_\alpha \times \{\alpha\}$. ($V_\alpha$ being a level of the usual hierarchy of sets).

Define $(x, \alpha) \in_{TTTU} (y, \beta)$ as $\alpha < \beta \wedge x \in y \wedge y \in V_{\alpha+1}$. When type β is being considered relative to its extensions over type α, the elements of $(V_\beta \setminus V_{\alpha+1}) \times \{\beta\}$ are treated as non-sets: $\mathtt{set}^\alpha_{\mathrm{TTTU}}(x, \beta)$ is defined as $x \in V_{\alpha+1}$ for $\beta > \alpha$.

Further, this shows that NFU is clearly consistent with Choice and with the negation of Infinity. Mathematics in NF is limited by the failure of choice (and lack of knowledge about

what choice-like principles may be safe to assume); mathematics in NFU is quite standard, though it may look unfamiliar because of the attention to stratification which is required.

We note here as we have noted elsewhere that Jensen's result completely vindicates Quine's foundational program. Specker's results showed that the passage from assuming the ambiguity scheme suggested by the symmetry of the theory to flat identification of the types could be justified. Jensen's argument shows that on the mathematically harmless assumption of urelements, the assumption of the ambiguity scheme can itself be justified.

The fact that a natural implementation of the Quine program leads to existence of urelements is perhaps startling, and the mathematical problem of whether we can get strong extensionality in NFU (and so get NF) remains technically interesting (and hard). But it has no philosophical interest: all philosophical charms of NF (that it has a universal set and complements, Frege natural numbers, and Russell-Whitehead cardinals and ordinals) are already found in NFU.

This framework for presenting the relationship between TTTU + Amb and NFU allows a straightforward account of why we can suppose that the same sentences hold in the model of TTTU+Amb and in the model of NFU.

Add to our language an external well-ordering $\leq_0$ of the model of TTTU + Amb we start with. Add to our language the relation $\leq=$ $\{((x, \alpha), (y, \beta)) : \alpha = \beta \wedge (x, \alpha) \leq_0 (y, \beta)\}$. Note that we cannot expect comprehension axioms for formulas containing $\leq$. Notice that we can define a Hilbert symbol $(\theta x : \phi(x))$ as referring to the first item a in the type of x such that $\phi(x)$ is true, if there is one, and otherwise the first object in the type. Apply the construction above to get consistency of TTTU + Amb with $\leq$ allowed to occur in instances of Amb, which provides a description of its own term model made up of Hilbert symbols. Then identify Hilbert symbols differing only by a shift of types to obtain a model of NFU in which the same sentences not mentioning $\leq$ hold as in the original model of TTTU + Amb.

14

Passing to the question of Con(NF)

NF is an extension of NFU, so the results above apply to NF as well. Given a model of TTT = TTTU + versions of strong extensionality for every sequence of types, we can construct a model of NF; in fact, the existence of a model of TTT is precisely equivalent to the existence of a model of NF.

The difficulty is that while TTTU is rather weird, on reflection it has quite natural models. This does not seem to be the case for TTT. Each type in TTTU is interpreted as the power set of *each* type below…plus some urelements. And this can be managed easily with the help of the cumulative hierarchy. Each type in TTT must be interpreted in effect as the power set of each lower type. This means that the power sets in question cannot be honest power sets, by simple considerations of cardinality.

In the remaining slides I will try to give the description of a model of TTT, briefly, without proving that it has stated properties. The proofs are nasty, and the coding techniques going into building the model are nasty.

Choose a limit ordinal λ. Ordinals below λ will be type indices.

Choose a regular uncountable cardinal $\kappa > \lambda$. Sets of size $< \kappa$ are called small, all others are called large.

Provide a set we call type -1 of a large cardinality μ (which will be the common cardinality of all the types from the standpoint of the metatheory). We stipulate that every element of type -1 is an ordered triple with first component -1; the equivalence relation $=_{TTT}$ which represents equality in our model agrees with equality on type -1. The size μ of type

-1 is a strong limit cardinal of cofinality at least κ.

Type -1 will be partitioned into disjoint sets of size κ called "litters". Each litter L has an associated "local cardinal" $[L]$, the set of all subsets of type -1 with small symmetric difference from L. Elements of local cardinals are called near-litters, and if $N \in [L]$ we also say that $[L]$ is the local cardinal of N and may write it $[N]$.

A code for a set of type $\alpha > -1$ is a triple (α, β, B) in which $-1 \leq \beta < \alpha$ and B is a set of codes for sets of type β. There are restrictions on what sets B can appear in codes, which will unfold.

The first restriction is that each element of B is a representative code taken from an equivalence class under $=_{TTT}$. The representative codes will be described below.

We postulate that $(\beta, \gamma, G) \in_{TTT} (\alpha, \beta, B)$ iff $(\beta, \gamma, G) \in B$. We can further assert that

$$(\alpha, \beta, B) =_{TTT} (\alpha, \beta, B')$$

iff $B = B'$: this follows from the first stipulation and the extensionality of TTT directly.

We expect though that (α, β, B) has an extension G of type $\gamma (\neq -1)$ elements for $\beta \neq \gamma < \alpha$, and we expect $(\alpha, \gamma, G) =_{TTT} (\alpha, \beta, B)$. We will then have $(\gamma, \delta, D) \in_{TTT} (\alpha, \beta, B)$ iff $(\gamma, \delta, D) \in G$ (where $.(\gamma, \delta, D)$ is representative).

We can already stipulate that $(\alpha, \beta, \emptyset) =_{TTT} (\alpha, \gamma, \emptyset)$

We provide the scaffolding for our solution for alternative extensions of set codes. We provide injective maps $f_{\alpha, \beta}$ for each pair of distinct types from type α into the set of local cardinals. The ranges of $f_{\alpha, \beta}$ and $f_{\alpha, \gamma}$ are disjoint for $\beta \neq \gamma$. α can be -1 in $f_{\alpha, \beta}$; β cannot. We state the exact meaning of *injective* here: $f_{\alpha, \beta}(b_1) = f_{\alpha, \beta}(b_2)$ iff $b_1 =_{TTT} b_2$; it will be injective in the usual sense on the representative codes.

We define $A_\gamma(\alpha, \beta, B)$, where B is nonempty, as $(\alpha, \gamma, \{(\gamma, -1, N) : b \in B \wedge N \in f_{\beta,\gamma}(b)\})$. We stipulated that the set of codes used in the model (codes which have representatives) is closed under all A_γ's. Notice that -1 is not a possible value of γ here: there are codes with -1 extensions built in (used here!), but we do not attempt to assign every code a -1 extension.

purely motivational remark: The maps A_γ will be used to assign alternative extensions. Note that each object will have an extension (presumably interesting and complicated) given in its representative code, and all of its other extensions will be unions of local cardinals in type -1, as embedded into other types. The key idea here is that we will use symmetries which turn type -1 and the subsets of type -1 into featureless junk, but in such a way that essentially arbitrary sets of sets of type

19

-1 objects (of the particular sort "union of local cardinals") can be added by constraining the permutations defining the symmetry suitably. This kind of "formless junk" plays the role that the urelements play in the construction of a model of TTTU. But making junk sets is much harder than making urelements.

We define A^{-1} as the union of the inverses of all the maps A_γ, which will be a (very partial!) function, because of the disjointness of ranges of functions $f_{\beta,\gamma}$ and $f_{\beta,\delta}$ when $\gamma \neq \delta$.

The idea is that for a representative code (α, β, B) (initially, one with B nonempty and which is not an image under A_β), its extension in type $\gamma \neq \beta$ is the γ extension of $A_\gamma((\alpha, \beta, B))$, with which code it is equivalent under $=_{TTT}$.

Now of course $A_\gamma((\alpha, \beta, B))$ cannot be representative: its β-extension is the β extension of (α, β, B) and its δ-extension for some third δ is the δ-extension of $A_\delta((\alpha, \beta, B))$

Now any code with two iterated images under A^{-1} is orphaned, so we treat these codes as representative, and in general the parity of the number of iterated images under A^{-1} of a code

determines whether the code is a representative code.

Choosing the maps $f_{\alpha,\beta}$ cleverly ensures that all codes have finitely many iterated images under A^{-1} so this scheme works. Impose a well-ordering on each type α. An element of type α is mapped by $f_{\alpha,\beta}$ to a range of local cardinals whose near-litter elements can be packaged as elements of type β by assigning them as type -1 extensions. Ensure that an element of type α is mapped to elements of type β which all appear in the order on type β at an ordinal position larger than the ordinal position in type α at which the original element of type α occurs. This can be arranged even if $\beta > \alpha$: we can have the ranges of the $f_{\alpha,\beta}$'s set from the beginning of the construction, and the ordinal positions for elements of each type with near-litter -1-extensions (we know what these extensions are at the outset) reserved at the outset. This condition ensures that no code can have infinitely many iterated images under A^{-1}.

We obviously must be intending to impose a restriction on what third projections set codes can have, as we appear to expect all the types to be the same size as type -1 in order to be able to construct the maps $f_{\alpha,\beta}$.

The restriction is a matter of symmetry. This is roughly speaking a Frankel-Mostowski construction (the fact that we are conducting it in a typed structure makes it look different).

We stipulate that in all cases $(\alpha, \beta, \{b\})$ is a set code for $b \in B$ a representative code (we provide for singletons).

A -1-allowable permutation is a permutation of type -1 whose elementwise action sends local cardinals to local cardinals: the elementwise image of any litter must be a near-litter.

We define an allowable permutation of type $\alpha \neq -1$ as a permutation of type α which sends (α, β, B) to $(\alpha, \beta, \pi_\beta "B)$ where π_β is an allowable permutation of type β such that

$$\pi(\alpha, \beta, \{b\}) = \pi(\alpha, \beta, \{\pi_\beta(b)\}).$$

Note that given π we can for each $\beta < \alpha$ determine what π_β must be and then check whether it is a β-allowable permutation: we are assuming in defining α-allowable permutations that

we already know what β-allowable permutations are for $\beta < \alpha$.

There is a coherence condition with $=_{TTT}$: an α-allowable permutation must respect this relation. Any $(\alpha, \beta, \{b\})$ is not in the domain of A^{-1} and so is equivalent to (for any γ distinct from β and less than α) $A_\gamma(\alpha, \beta, \{b\})) = (\alpha, \gamma, \{(\gamma, -1, N) : N \in f_{\beta,\gamma}(b)\})$. So the coherence condition tells us that $(\alpha, \beta, \{\pi_\beta(b)\})$ must be equivalent to $(\alpha, \gamma, \pi_\gamma ``\{(\gamma, -1, N) : N \in f_{\beta,\gamma}(b)\}) = (\alpha, \gamma, \{(\gamma, -1, (\pi_\gamma)_{-1} ``N) : N \in f_{\beta,\gamma}(b)\})$, from which we can see (take my word for it?) that what is required is that $f_{\beta,\gamma}(\pi_\beta(b)) = [(\pi_\gamma)_{-1} ``N]$ for any $N \in f_{\beta,\gamma}(b)$. Some notation chasing reveals that this exactly captures the coherence condition.

It is straightforward to check that allowable permutations send representative codes to representative codes.

For any nonempty finite subset A of $\lambda \cup \{-1\}$, with largest element α, and α-allowable permutation π, define $\pi_{\{\alpha\}}$ as π and define π_A as $(\pi_{A \setminus \{\min(A)\}})_{\min(A)}$ if A has at least two elements. We get a whole suite of permutations of lower types induced by an allowable permutation of type α along different paths through the types (and one should note that these permutations are chosen fairly freely of each other; the coherence condition is not a strong constraint).

A α-support set is a small (size $< \kappa$) set of pairs (x, A) where the maximum of A is α and the minimum of A is the type of x, and the extension of x as a representative code is a subset of type -1, either a singleton or a near-litter. An object X of type α has support S if any allowable permutation π such that π_A fixes x for each $(x, A) \in S$ has $\pi(X) = X$.

We stipulate that the elements of type α are exactly the set codes with α-supports. Notice that we can apply any allowable permutation to all set codes with first term α before we have actually determined the extent of type α, since evaluation reduces to information about permutations of lower types which we may suppose we have already worked out.

And that is it! The model is described. But showing that it works is horrible.

This hinges on properties of what is clearly a Frankel-Mostowski sort of construction over type -1. I can tell you at once (and if I had more time I could prove in a lecture like this) that the symmetric sets of type -1 objects (in any type) are sets with small symmetric difference from small or co-small unions of litters (and externally there are only μ of these).

The fun thing about this FM construction is that one can impose further conditions (as I do in admittedly confusing ways in this construction) which preserve exactly what the sets of type -1 objects are but add arbitrary new unions of local cardinals (lots of sets of sets of type -1 objects). That is what the maps $f_{\alpha,\beta}$ are doing in the construction: we are in effect creating copies of each lower type as unions of local cardinals in each higher type (and then doing further things to them to make everything work properly for TTT).

It is straightforward (yes, really) to prove (under assumptions that other things work correctly) that predicative TST in a tangled version holds in this structure. Predicative TST doesn't allow us to mention higher types in the definition of a set in a given type.

There are two harder things to prove.

You get from predicative TST to TST by adding the axiom of set union. We need to show that union holds in this structure, for which it is sufficient to show that any set of singletons of type β objects in type α has an α-support which is in fact simply a translate of a β-support: in effect, we dont get new subsets of lower types when we impose more symmetry to get the desired conditions on higher types.

We need to verify that all the types are actually the same size. For this (and for the

other results, this is a universal tool), we need to show that the allowable permutations act rather freely. There is a precise statement of the degree of freedom of action needed, which itself requires more than one slide to set up properly.

A note on forms of this proof, if we get to it

There are numerous versions of this proof, which fall into two main families.

The other family takes a different approach (my original one), which involves an FM argument of the usual kind showing the existence of a certain kind of system of cardinals in a model of ZFA without choice.

A tangled web of cardinals is a map τ from nonempty finite subsets of a limit ordinal λ to cardinals such that

1. For each A with $|A| \geq 3$, $2^{\tau(A)} = \tau_{A \setminus \{\mathtt{min}(A)\}}$. (Use the Scott definition of cardinals since choice is not assumed and the $\tau(A)$'s cannot be alephs).

2. For each A with $|A| > n$, the theory of the first n types of the natural model of TST with base type of cardinality $\tau(B)$ is the same as the theory of the natural model of TST with base type of cardinality $\tau(A)$ if the smallest n elements of B are the same as the smallest n elements of A. (The natural model of TST with base type X has as type i a copy of the full ith iterated power set of X).

It is a theorem (proved in a way reminiscent of Jensen's argument for Con(NFU): we do not give details here) that if there is a model of ZFA with a tangled web of cardinals, NF is consistent. The proof is simpler in the sense that one works in ZFA with families of atoms intended to have the cardinalities $\tau(A)$. But getting the second condition on the tangled web, with its metamathematical flavor, involves details which at bottom seem to end up building a model of tangled type theory anyway.