The NF consistency problem

M. Randall Holmes
Boise State University

June 5, 2019

Abstract

In two successive one hour talks, I will give a brief account of the NF consistency problem, explain why consistency of NF follows from the existence of something called a ”tangled web of cardinals” in a model of ZFA, and then give the barest hint of a Frankel-Mostowski construction of a model of ZFA in which there is a tangled web. The third goal is difficult and the degree of success to be expected should not be overrated.

This version posted after the actual talks contains typo corrections and some additional comments.

We begin with a theory which we call TST, the simple typed theory of sets.

TST is a first-order theory with sorts indexed by the natural numbers. We manage sorts by providing a function $\texttt{sort}(`x')$ taking variables to natural numbers. Its primitive predicates are equality and membership, subject to the constraints that '$u = v$' is well-formed iff $\texttt{sort}(`u') = \texttt{sort}(`v')$ and '$u \in v$' is well-formed iff $\texttt{sort}(`u') + 1 = \texttt{sort}(`v')$.

The axiom (schemes) of TST are

1. an axiom scheme of extensionality: each sentence

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

   which is well-formed is an axiom.

2. an axiom scheme of comprehenson: each sentence $(\exists A : (\forall x : x \in A \leftrightarrow \phi))$ which is well-formed and in which A does not occur in the subformula ϕ is an axiom.

   We refer to the witness to this instance of comprehension (unique by extensionality) as $\{x : \phi\}$.

The resemblance of these axioms to those of "naive set theory" is not accidental.

We note the existence of the subtheory TST_n in which the indices of sorts are restricted to natural numbers less than n.

We define a *natural model* of TST_n with base cardinal κ as a model of TST_n in which type 0 is implemented by a set of cardinality κ and each type $i+1$ ($i \leq n-2$) is the power set of the set implementing type i. It should be clear that the model is determined up to isomorphism by n and κ.

The theory above appears remarkably late. There are hints of this idea in Russell, Principles of Mathematics, 1904(?). The type system of *Principia Mathematica* is far more complicated. The possibility of a simple type theory of this kind was seen by Norbert Wiener when he first pointed out that an ordered pair can be defined in set theory (1914). The first formal description of this exact theory (though written in a way which makes determining that it is the same theory very difficult) seems to be by Tarski in 1929(?). Hao Wang has a nice discussion of this somewhere.

This theory has a very high degree of symmetry for reasons already noted by Russell in his development using the much more complicated type system of Principia. With this much simpler type system, the scope of what Russell called "systematic ambiguity" is greater.

Let `raise` be a bijective map from variables to variables with positive sort such that
$\mathtt{type}(\mathtt{raise}(`x'))=\mathtt{type}(`x')+1$ in all cases.

For any formula ϕ, define ϕ^+ as the result of replacing every variable in ϕ with its image under `raise`. Note that for standard logical reasons the effect of this operation on sentences (formulas with no free variables) does not depend at all on the details of `raise`.

Now observe that it is evident that if we can prove ϕ, we can prove ϕ^+, and if we define a mathematical construction using a set abstract $\{x : \phi\}$, there is an exactly analogous construction $\{x^+ : \phi^+\}$ producing an object one type higher. (x^+ being a convenient way of writing the type-raised version of x).

It is not accurate to say that we can prove the same things about types 1,2,3... that we can prove about types 0,1,2,...; in fact, we know more about the former because we know more about type 1 than we know about type 0. But these facts, and the inconvenience of reproducing each theorem and construction at each new type (already noted in effect by Russell) let Quine to propose in 1937 that perhaps the types should be taken to be all the same.

NF is a first order single sorted theory with equality and membership as its primitive logical relations.

The axiom (schemes) of TST are

1. an axiom of extensionality:

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

   is an axiom.

2. an axiom scheme of comprehension: each sentence $(\exists A : (\forall x : x \in A \leftrightarrow \phi))$ in which A does not occur in the subformula ϕ and in which there is an assignment of sorts to the variables appearing in ϕ which would make ϕ a well-formed formula of TST is an axiom.

   We refer to the witness to this instance of comprehension (unique by extensionality) as $\{x : \phi\}$.

We devote a slide to showing that the reference to the language of TST in the previous slide is unnecessary. We can say instead that in an instance of comprehension, the formula ϕ must have an associated function σ mapping variables appearing in ϕ to natural numbers with the properties that for any subformula '$u = v$' of ϕ it will be the case that $\sigma(`u') = \sigma(`v')$ and for any subformula '$u \in v$' of ϕ, it will be the case that $\sigma(`u')+1 = \sigma`(v')$. The function σ is called a "stratification" of ϕ and a formula with a stratification is said to be "stratified".

Even this is not needed. The stratified comprehension scheme is equivalent to a finite conjunction of its instances, so it can be stated without any reference to even relative types.

Specker showed in 1962 that Quine's intuition in formulating the theory was correct.

Define the Ambiguity Scheme as the set of closed sentences of the form $\phi \leftrightarrow \phi^+$.

Specker proved (among other things) that NF is precisely equiconsistent with TST + the Ambiguity Scheme. He proved more, but this is what we need here.

We explain the sense in which we think this justifies Quine. Quine noted that if we can prove ϕ, we can also prove ϕ^+. From this he formed the daring conjecture that it is reasonable to suppose that if ϕ is true, ϕ^+ is true ($\phi \leftrightarrow \phi^+$ would follow from this by applying the same principle to $\neg\phi$ as well). Specker hasn't justified this with his 1962 results: what he has justified is Quine's passage from successive types satisfying analogous sentences to successive types being precisely the same domain.

The world of NF has certain peculiar characteristics which some find philosophically appealing. The universe is a set and all sets have complements. Cardinal numbers can be taken to be Frege cardinals (3 is the set of all sets with three elements) and ordinals can be taken to be Russell-Whitehead ordinals.

The paradoxes are evaded in interesting and exciting ways.

I am not going to talk about any of this because it is entirely irrelevant to the NF consistency problem.

The reason that the presence of big sets in NF is irrelevant IS relevant to our project, so I will explain. In 1969, Jensen showed that the system NFU obtained by weakening extensionality to allow many urelements (objects with no elements) is consistent. NFU allows all the same constructions of big sets, and avoids the paradoxes in the same exciting and interesting ways, and it can be seen in well-understood models in the usual set theory why all of this works.

I have claimed elsewhere that any philosophical interest which NF has is equally well served (in fact in my view better served, for reasons mentioned below in passing) by NFU.

An objection which some have raised that NFU is unsatisfactory because not a theory of pure sets seems to me misguided for two reasons. Nothing but custom makes us think that restricting to pure sets has philosophical merit (it is admittedly mathematically convenient). The world of even mathematical experience contains a lot of things which may reasonably be supposed not to be sets. And restricting to pure sets is a possible maneuver in set theory in the style of Zermelo, but an impossible or at least difficult one in a set theory driven by stratified comprehension, because the predicate of being a pure set is unstratified in a rather horrible way.

We prove the consistency of NFU.

Let X be a sequence of sets indexed by a limit ordinal λ such that $\mathcal{P}(X_\alpha) \subseteq X_\beta$ for each $\alpha < \beta < \lambda$. Such a sequence of sets would be obtained if $X_\alpha = V_\alpha$ for each $\alpha < \lambda$, for example. It is worth noting that with some indirection it is enough to assume that $|\mathcal{P}(X_\alpha)| < |X_\beta|$ for each $\alpha < \beta$, but writing the proof in this style would multiply indices with no corresponding gain in intelligibility.

Let s be any strictly increasing sequence of elements of λ. We present an interpretation of the language of TSTU (TST with extensionality weakened to apply only to objects with elements) using the sequence X. Each type i is interpreted as $X_{s(i)}$. Equality is interpreted as usual. Where u is a type i variable referring to $x \in X_{s(i)}$ and v is a type $i+1$ variable referring to $y \in X_{s(i+1)}$, we define $u \in_{TSTU} v$ as holding iff $x \in y \wedge y \in \mathcal{P}(X_{s(i)})$. Note that we are reinterpreting all elements of $X_{s(i+1)} \setminus \mathcal{P}(X_{s(i)})$ as urelements.

Let Σ be a finite set of formulas of the language of TSTU (which might be the same as that of TST, but often includes a constant for the empty *set* in each type as a formal convenience). Let n be chosen so that Σ is a finite set of formulas of the language of TSTU_n. Define a partition of λ determined for each n-element subset A of λ by the (common) assignment of truth values to formulas in Σ holdng in interpretations of TSTU defined as above with $\texttt{rng}(s\lceil n) = A$. This partition has an infinite homogeneous set H by Ramsey's theorem. Choose a strictly increasing sequence h in H. In the interpretation of TSTU determined by h, the Ambiguity Scheme holds for formulas $\phi \in \Sigma$. From this it follows that the entire Ambiguity Scheme is consistent by compactness, so TSTU + Ambiguity is consistent. Specker's 1962 methods apply to NFU just as they do to NF, so NFU is consistent.

It can further be noted that NFU does not prove Infinity (all the sets X_α could be finite) though any model of NFU is certainly externally infinite.

NFU is consistent with Infinity (the X_α's could be taken to be infinite) and with Choice (which would be inherited from the ambient set theory if it held there).

NFU is consistent with strong axioms of infinity expressible in its own terms, which can be arranged by letting λ be large enough to satisfy interesting strong partition properties which can translated into interesting unstratified axioms to adjoin to NFU. I can't help mentioning this because it is a lovely program, but only one such axiom might get passing reference here, Rosser's Axiom of Counting, which I will introduce if I need it.

An original contribution of ours in 1995 was to point out a way that Jensen's argument could be adapted to Con(NF), though this did not yield a consistency proof because the conditions required for the Jensen argument to go through looked wildly unlikely.

Note that in the argument above we are in effect working in a model of TSTU with types indexed by the ordinals below λ, and providing membership relations $\in_{\alpha,\beta}$ for type α objects in type β objects for each $\alpha < \beta < \lambda$. In this context, the fact that types with index $> \alpha + 1$ look too large to be "power types" of type α is actually not a problem at all: we just interpret the "too many" additional objects as urelements.

An extensional type theory TTT (for "tangled type theory") of this kind can be described, and shown equiconsistent to NF, though the existence of models of this theory is not obvious at all.

Let λ be a limit ordinal. TTT is a first order theory with sorts indexed by λ. A formula $u = v$ is well-formed iff u and v are of the same sort. A formula $u \in v$ is well-formed iff the sort index assigned to v is greater than the sort index assigned to u.

Let s be a strictly increasing sequence of ordnals below λ. Let ϕ be a formula of the language of TST: define ϕ^s as the result of replacing each TST variable of type i in ϕ with a TTT variable of type $s(i)$ in an injective way. The axioms of TTT are precisely the sentences ϕ^s where ϕ is an axiom of TST.

Each element of a type β has an extension as a subset of each lower type α, and each of these extensions uniquely determines it. Sets can be defined as in TST on each increasing path of types: the relationship between sets defined on different paths is altogether unclear.

Cantor's theorem tells us at once that we cannot construe all of these "power sets" as true power sets. Consideration of the cardinality of three types in a model will rule this out.

NF is consistent iff TTT is consistent.

Given a model of NF, use that very model (or of you prefer disjoint copies of the model) to implement each type of a model of TTT, and use the membership of the model to determine the membership between any two types.

Given a model of TTT, let Σ be a finite set of formulas of the language of TST. Let n be chosen so that Σ is a finite set of formulas of the language of TST_n. Define a partition of λ determined for each n-element subset A of λ by the (common) assignment of truth values to formulas ϕ^s for ϕ in Σ where $\mathtt{rng}(s\lceil n) = A$. This partition has an infinite homogeneous set H by Ramsey's theorem. Choose a strictly increasing sequence h in H. In the interpretation of TST determined by h, the Ambiguity Scheme holds for formulas $\phi \in \Sigma$. From this it follows that the entire Ambiguity Scheme is consistent by compactness, so TST + Ambiguity is consistent. Specker's 1962 methods show that NF is consistent.

I provide a single slide to report much pain. TTT is impossible to work in. It is wildly counterintuitive and the opportunities for conceptual and notational mistakes are endless.

Also in our 1995 paper, we determined how to reduce the problem of consistency of NF to the existence of certain patterns of cardinals in Zermelo-style set theory.

We work in ZFA (ZF with atoms). We won't need atoms in this definition, but we do need them in our actual work below.

Let λ be a limit ordinal. A *tangled web* of order λ is a function τ from nonempty finite subsets of λ to cardinals with the following properties:

1. For each A with $|A| \geq 2$,
$$2^{\tau(A)} = \tau(A \setminus \{\mathtt{min}(A)\}).$$

2. If $|A| \geq n$, the theory of the natural models of TST_n with base cardinal $\tau(A)$ is precisely determined by the n smallest elements of A.

The existence of a tangled web implies Con(NF).

Given a tangled web τ, let Σ be a finite set of formulas of the language of TST. Let n be chosen so that Σ is a finite set of formulas of the language of TST_n. Define a partition of λ determined for each n-element subset A of λ by the truth value of each $\phi \in \Sigma$ in the natural models of TST_n with base cardinal $\tau(A)$. This partition has a homogeneous set B of size $n+1$ by Ramsey's theorem. The natural model of TST_{n+1} with base cardinal $\tau(B)$ has type 1 of cardinality $2^{\tau(B)} = \tau(B \setminus \{\mathtt{min}(B)\})$. The theories of its two submodels of TST_n are the same because one is determined by the n smallest elements of B and one is determined by the n largest elements, and these belong to the same compartment of the partition and so determine the same theory, and so we see that Ambiguity restricted to Σ is consistent with TST_{n+1} (and actually with TST because n can be taken as large as desired). It follows that TST + Ambiguity and NF are consistent.

The underlying idea of a tangled web is to undo the notion of identifying elementarily equivalent structures which underlies Quine's intuition for NF just enough to get rid of the impossible power sets of a model of TTT. Replace each type α of a model of TTT with copies indexed by each finite subset A of λ with α as its minimum element, with the idea that the power set of type A will be type $A \backslash \{\mathtt{min}(A)\}$ and finite increasing sequences of types in the model of TTT are taken to have the same theory as the natural models with base type the set of indices of the types in the sequence.

Our program is then to construct a model of ZFA in which there is a tangled web of cardinals.

A scandal

I have so far failed to say anything about a well-known fact about NF.

Specker showed in 1954 that NF disproves Choice and so proves Infinity.

I showed in my 1995 paper how to carry out a version of Specker's argument directly in TTT. Similarly, the existence of a tangled web contradicts choice.

This fact suggests the strategy that actually succeeds: use the Fraenkel-Mostowski construction of models of ZFA in which choice fails to construct a tangled web. This can be done, but it is not at all obvious how to do it, and there I was stalled for fifteen years.

The salient feature of a tangled web (inherited from its conceptual origins in the description of tangled type theory) is that natural models of initial segments of type theory which "look the same" (are elementarily equivalent) may have power sets of their top types and so extensions to a longer initial segment of type theory which look very different. The key idea of the next part of the argument is a technique for producing models of initial segments of type theory in which one has an enormous degree of freedom in determining what the next power set looks like.

Let κ be a regular uncountable cardinal, fixed for the rest of the talk. We refer to sets with cardinality $< \kappa$ as small and all other sets as large.

Let μ be a strong limit cardinal greater than κ with cofinality at least κ, fixed for the duration of the talk.

We describe something called a "clan" which is a basic building block of our argument.

A clan is a collection K of μ atoms with a partition $\Lambda(K)$ into collections of size κ which we refer to as "litters". We further define for each $L \in \Lambda(K)$ its "local cardinal" $[L]$, which is the collection of subsets of K with small symmetric difference from L. We define $\Pi(K)$ as $\{[L] : L \in \Lambda(K)\}$, the collection of all local cardinals of litters included in K.

The permutation group of our FM model will consist of permutations which fix $\Pi(K)$ [and fix $\Pi(K')$ for each clan K' present] and further fix a map P_K, the parent map for K, chosen in the course of the construction. P_K is a bijection from a subset of $\Pi(K)$ to some possibly quite remote part of the model.

We are given a indexing of the clans we are using (there are $\leq \mu$ of them). We define a sequence M_α of FM models: each M_α has as its permutation group G_α the collection of permutations of the atoms whose action fixes every Π_{K_β} and further fixes each P_{K_γ} for $\gamma < \alpha$.*

A near-litter included in a clan K is an element of $\bigcup \Pi(K)$, a subset of K with small symmetric difference from a litter. If N is a near-litter, it has small symmetric difference from a uniquely determined litter N°.

*It is probably helpful to note that in the paper this is based on what is here written P_{K_α} is written P_α.

A support is a well-ordering $\leq_S$ with domain S, a small set of atoms and near-litters, with the property that distinct near-litter elements of S are disjoint. An object X has α-support $\leq_S$ iff every permutation in G_α whose action fixes $\leq_S$ also fixes X. A set is α-symmetric iff it has an α-support. The elements of M_α are the sets which are hereditarily α-symmetric, and standard techniques show that M_α is a model of ZFA.

Note that for $\alpha > \beta$, $M_\beta \subseteq M_\alpha$, since any β-support is clearly a α-support, as $G_\alpha \subseteq G_\beta$.

The basic idea is that at stage $\alpha + 1$ we add more subsets of $\Pi(K_\alpha)$ by requiring P_{K_α} to be a set, making these sets the same size as sets defined at an earlier stage of the construction somewhere else in the model.

Conditions which must be satisfied by P_{K_α}: the range of P_{K_α} must have empty α-support (it must be invariant under all permutations in G_α).

Further, we provide in advance a partition of each $\Pi(K)$ into κ disjoint "echelons" $\Pi_\alpha(K)$ of cardinality μ for $\alpha < \kappa$ or $\alpha = -1$. The echelon of a local cardinal is the α such that the local cardinal belongs to $P_\alpha(K)$. A litter has the echelon of the local cardinal to which it belongs. An atom has the echelon of the litter to which it belongs. A near-litter N has echelon the supremum of the echelons of N° and the elements of $N \Delta N^\circ$.

A value $P_K(x)$ must have a support all of whose elements have echelon less than that of x. If x in $\Pi(K)$ has echelon -1, it cannot be in the domain of P_K.

Analysis of the action of permutations requires more detail in the notion of support. An α-strong support is a support in which

1. each atom is preceded by the litter containing it

2. each near-litter in the domain is a litter

3. each litter L included in K_β ($\beta < \alpha$) belonging to the range of the support is preceded in the support by the elements of a β-support of $P_{K_\beta}([L])$, if this is defined, in which all elements have echelon less than the echelon of L.

Strong supports have numerous uses in the mechanics of the argument, which we will have little or no opportunity to discuss in detail.

A remark on terminology

I might refer to $P_K([L])$ as the parent of $[L]$ or indeed of L, of a near litter N with $N^\circ = L$ or of an atom in L.

The notation P might be used for the union of the P_K's.

The set P_K"$(\Pi(K))$ might be called the parent set of K.

Freedom of actions of the permutations

Define a α-local bijection as an injective map with the same domain and range, and with range a set of atoms with small intersection with each litter (empty is a case of small) and all local cardinals which are not in the domain of some P_{K_β} for $\beta < \alpha$.

Note that permutations in our groups G_β send litters to near-litters (because they fix each $\Pi(K)$); an exception of a permutation π is an atom x in a litter L which is mapped out of $\pi(L)^\circ$ by π or out of $\pi^{-1}(L)^\circ$ by π^{-1}. There are a small set of exceptions of a permutation in each litter.

The Freedom of Action Theorem states that any α-local bijection can be extended to a permutation in G_α with no exceptions other than elements of the domain of the local bijection.

We do not present the proof in detail (we might discuss it briefly live) except to note that it relies on induction on strong supports.

It is a straightforward consequence of the Freedom of Action theorem that the power set of a clan K in any of the M_α's is the collection of subsets of K in the ground interpretation of ZFAC which have small symmetric difference from a small or co-small union of litters.

In particular, the cardinalities of litters are κ-amorphous. I might outline the proof of this result as an illustration of why strong supports are useful.

Note that any small subset of the domain of an M_α belongs to that M_α: a small collection of strong supports of elements can be reorganized into a single strong support for the small set.

The results so far are all based an abstract scheme for introducing subsets of $\Pi(K)$'s. Notice that we already have in each clan and its power set an example of elementarily equivalent (indeed, externally isomorphic) initial sequences of natural models of the theory of types in which the next power set may be wildly different in different examples: for each set K, the requirement that P_K be fixed by the permutations in our groups has effects different for each K on what sets we find in the double power set of K. Further, if a clan K' is found in the image of P_K this will add a nondescript set to $\mathcal{P}^2(K)$ and quite unexpected contents of $\mathcal{P}^2(K')$ to $\mathcal{P}^3(K)$. This process can be arranged for iterated power sets of higher index.

Now we describe the actual functions P_K which we use, thus giving a complete description of our FM models (the limit model being the model in which there is a tangled web).

Let λ be a limit ordinal $\leq \kappa$ ($\leq \mathtt{cf}(\mu)$ probably suffices).

We define a well-ordering $\ll$ on finite subsets of λ:

1. $A \ll \emptyset$ for every nonempty A.

2. if $\mathtt{max}(A) < \mathtt{max}(B)$ then $A \ll B$.

3. if $\mathtt{max}(A) = \mathtt{max}(B)$ then

$$A \ll B \leftrightarrow A \setminus \{\mathtt{max}(A)\} \ll B \setminus \{\mathtt{max}(B)\}.$$

That these conditions indeed determine a unique well-ordering of the finite subsets of λ I leave as an exercise.

For each nonempty subset A of λ, define A_1 as $A \setminus \{\mathtt{min}(A)\}$ and define A_0 as A and A_{n+1} as $(A_n)_1$ where this is defined.

The order type of our sequence of K_α's is $\mathtt{ot}(\ll) + 1$.

For each finite subset A of λ, we define $\mathtt{clan}[A]$ as the K_α such that α is the order type of the restriction of $\ll$ to sets $\ll A$. We use the notation $\mathtt{ind}(\mathtt{clan}[A])$ for this α.

For each nonempty A, we define $P_{\mathtt{clan}[A]}$ as a bijection from a subset of $\Pi(\mathtt{clan}[A])$ to

$$(\mathtt{clan}[A_1] \cup \bigcup_{\beta<\mathtt{min}(A)} \mathcal{P}^2(\mathtt{clan}[A \cup \{\beta\}])) \cap M_{\mathtt{ind}(\mathtt{clan}[A])}.$$

We define $P_{\mathtt{clan}}[\emptyset]$ as a bijection from $\Pi(\mathtt{clan}[\emptyset])$ to μ. The model $M_{\mathtt{ot}(\ll)+1}$ is then constructed in all details that matter (it has a trailing clan of which we make no use).

This is admittedly confusing and may appear preposterous. Reasons why it isn't preposteroous and reasons why it does what we want will be outlined.

It is important to note that downward extensions of finite subsets of λ appear before those sets in our curious order.

There are two facts to be shown to assure ourselves that this is not preposterous. Neither has a proof of which I can give too much of a hint in the scope of this talk.

Absoluteness of iterated power sets: the first fact.

$$\mathcal{P}^2(\texttt{clan}[A \cup \{\beta\}]) \cap M_{\texttt{ind}(\texttt{clan}[A])}$$
$$= \mathcal{P}^2(\texttt{clan}[A \cup \{\beta\}]) \cap M_\gamma$$

for any $\gamma \geq \texttt{ind}(\texttt{clan}[A]$.

The approximations to double power sets appearing in the indicated model are actually the double power sets of every succeeding FM model we construct: no more sets are ever added.

More generally, if $|A| \geq n$,

$$\mathcal{P}^{n+1}(\texttt{clan}[A]) \cap M_{\texttt{ind}(\texttt{clan}[A_n])}$$

$$= \mathcal{P}^{n+1}(\texttt{clan}[A]) \cap M_\gamma$$

for any $\gamma \geq \texttt{ind}(\texttt{clan}[A_n]$.

and it is further an important feature of the proof that the restriction of the strong support of an element of $\mathcal{P}^{n+1}(\texttt{clan}[A \cup \{\beta\}])$ to atoms and litters belonging to or included in clans downward extending A_n is a support (and almost a strong support) for that element.

Size of iterated power sets of clans: the second fact

$$|\mathcal{P}^{n+1}(\mathtt{clan}[A]) \cap M_{\mathtt{ind}(\mathtt{clan}[A_n])}| = \mu \text{ if } |A| \leq n.$$

In the case $n = 1$, this shows that the purported range of $P_{\mathtt{clan}[A]}$ is actually small enough to be the image of a subset of $\Pi(\mathtt{clan}[A])$.

The proof is by an elaborate scheme of counting orbits, driven by an analysis of strong supports...*

*It is important to note, as brought out by an intelligent question at the actual talk, that the fact that μ is a strong limit cardinal comes into play right here, bringing the necessary consistency strength into the argument. One analyzes orbits of sets in an iterated power set of a clan by considering sets of orbits of their potential elements. The fact that there are not too many sets of orbits one type down depends (among many other things!) on μ being strong limit.

Elementary equivalence

Fix a finite subset A of λ. There is a map χ_A in the ground interpretation which sends each $\mathtt{clan}[B]$ to $\mathtt{clan}[B \cup A]$, where each element of B is dominated by each element of A, and commutes with each P_{K_β} for $\beta < \mathtt{ind}(\mathtt{clan}[A])$, when understood to be extended to sets in the obvious way. It is constructed by recursion on information in strong supports!

The map χ_{A_n} gives an external isomorphism between the initial segment of a natural model of type theory with $\mathcal{P}^{n+1}(\mathtt{clan}[B])$ as its top type and the the initial model of a natural model of type theory with $\mathcal{P}^{n+1}(\mathtt{clan}[B \cup A])$ as its top type, which is not visible in our FM model, but the fact that these natural models have the same first order theory is visible.

We prove a useful lemma about cardinalities of double power sets of clans in our final FM model.

Definition: We introduce the notation $\mathcal{P}_*(X)$ for the power set in the sense of our final FM model of an element X of that model. We introduce the notation $|X|_*$ for the cardinality of the set X in our final FM model in the sense of that model. We allow the usual overloading of $\leq$ to represent the order on cardinals in our model.

Definition: We define $\tau(A)$ as $|\mathcal{P}^2_*(\mathtt{clan}[A])|_*$.

Observe that

$$|P\,``(\Pi(\texttt{clan}[A])|_* \leq |\mathcal{P}^2_*(\texttt{clan}[A])|_*$$

is obvious (because $\Pi(K)$ is a subset of $\mathcal{P}^2(K)$ and the P_K's are all bijections and sets in the FM model). In fact, the stronger assertion

$$|\mathcal{P}_*(P\,``(\Pi(\texttt{clan}[A])))|_* \leq |\mathcal{P}^2_*(\texttt{clan}[A])|_*$$

is true, because the elements of $\Pi_*(\texttt{clan}[A])$ make up a pairwise disjoint family of elements of $\mathcal{P}^2_*(\texttt{clan}[A])$, so in fact

$$|\mathcal{P}_*(P\,``(\Pi(\texttt{clan}[A])))|_* \leq |\mathcal{P}^2_*(\texttt{clan}[A])|_*.$$

From this we get the corollaries that

$$|\mathcal{P}_*(\mathtt{clan}[A_1])|_* \leq |\mathcal{P}_*^2(\mathtt{clan}[A])|_*,$$

because $\mathtt{clan}[A_1]$ is included in the parent set of $\mathtt{clan}[A]$, and that

$$|\mathcal{P}_*^3(\mathtt{clan}[A \cup \{\beta\}])|_* \leq |\mathcal{P}_*^2(\mathtt{clan}[A])|_*,$$

where $\beta < \mathtt{min}(A)$, because

$$\mathcal{P}_*^2(\mathtt{clan}[A \cup \{\beta\}])$$

is included in the parent set of $\mathtt{clan}[A]$.

Double Power Set Theorem: Assume $|A| \geq 2$.

$$|\mathcal{P}_*^3(\mathtt{clan}[A])|_* = |\mathcal{P}_*^2(\mathtt{clan}[A_1])|_*.$$

That is, the FM model believes that

$$2^{\tau(A)} = \tau(A_1).$$

Proof: We have from above that

$$|\mathcal{P}_*^3(\mathtt{clan}[A])|_*$$

$$= |\mathcal{P}_*^3(\mathtt{clan}[A_1 \cup \{\mathtt{min}(A)\}])|_* \leq |\mathcal{P}_*^2(\mathtt{clan}[A_1])|_*$$

and that

$$|\mathcal{P}_*(\mathtt{clan}[A_1])|_* \leq |\mathcal{P}_*^2(\mathtt{clan}[A])|_*$$

whence

$$|\mathcal{P}_*^2(\mathtt{clan}[A_1])|_* \leq |\mathcal{P}_*^3(\mathtt{clan}[A])|_*,$$

from which the result follows.

The proof of this theorem presents the motivation for the strange details of our choice of parent sets in a nutshell.

And now we have a tangled web.

We define $\tau(A)$ as the cardinality of the double power set of $\texttt{clan}[A]$.

The fact that $2^{\tau(A)} = \tau(A_1)$ was just proved.

We need to show that the natural models with base type of cardinality $\tau(A)$ with $|A| \geq n$ have first order theory determined by the set $A \setminus A_n$ of the n smallest elements of A.

The model in question has lowest type $\mathcal{P}_*^2(\texttt{clan}[A])$ and highest type $\mathcal{P}_*^{n+1}(\texttt{clan}[A])$, and if $A \setminus A_n = B \setminus B_n$, then application of the inverse of χ_{A_n} followed by χ_{B_n} will give an external isomorphism from the model of TST$_n$ with base type the double power set of $\texttt{clan}[A]$ to the model of TST$_n$ with base type the double power set of $\texttt{clan}[B]$.

So both defining properties of a tangled web hold.