Consistency of Quine's NF

Murdoch J. Gabbay

29 March 2022

Thanks

Many thanks to LFCS and Edinburgh Informatics for the invitation to speak.

I'm here for you: if you don't follow then please just ask.

Why care about foundations of mathematics?

I probably don't need to push this case too hard at a *Laboratory for Foundations of Computer Science* seminar, but let me spell this out as I see it.

The study of the foundations of mathematics is not ivory tower maths. It's problem-solving — where the problem addressed is

> *What building blocks do we need to solve problems using rigorous mathematical thought?*

This is not an *abstract* question, so much as a *distilled* question. Consider . . .

Why care about foundations?

1. Theorem-provers = applied foundations.
   Lean, AGDA, COQ, Isabelle/HOL, and all others are explicitly implementations of foundations.
2. High-level programming languages = applied foundations.
   This is deliberately centre stage in e.g. Haskell, but is also visible in e.g. Python (think: lambda; iterators; class programming), or even in C (think: Turing machines).

Foundations are a way to study our relationship with our own understanding of what makes sense and is intuitive.

One strong intuition is that of 'a set' ...

Naive set theory

We carry an intuition of 'a set', as being a collection of things that we can add to and take away from.

Naive set theory makes this foundationally precise as follows:

1. Everything is a set.
2. If ϕ is a predicate in first-order logic (FOL) with $\in$, then the comprehension

$$\{a \mid \phi\} \quad \text{meaning} \quad \text{“the set of } a \text{ such that } \phi\text{”}$$

   is a set.

This is arguably the first, greatest, foundation. But ...

. . . naive set theory is inconsistent

Recall that famous inconsistency proof (Russell, 1902). Consider

$$R = \{a \mid a \notin a\}.$$

Then $R \in R \Leftrightarrow R \notin R$:

$$R \in R \quad \Leftrightarrow \quad R \in \{a \mid a \notin a\} \quad \Leftrightarrow \quad R \notin R.$$

Thus the system is inconsistent.

Much of 20th century foundational thought was devoted to escaping this inconsistency! Notably: ZFC, HOL, (dependent) types.

Quine's New Foundations / Typed Set Theory +

I may identify Quine's NF with the closely related system TST+, and write 'NF' and 'TST+' synonymously.

Quine proposed a system in 1937 which works like this:

1. Define levels to be numbers 0, 1, 2, ...
2. Everything is a set of some level.
3. If ϕ is a stratified predicate — we only form $b \in a$ when $\mathit{level}(a) = \mathit{level}(b) + 1$ — then the stratified comprehension $\{a \mid \phi\}$ is a set of level $\mathit{level}(a) + 1$.
4. Typical Ambiguity (TA): If ϕ is a *closed* predicate then $\phi \Leftrightarrow \phi^+$, where ϕ^+ is obtained by shifting every variable in ϕ up by 1.

Examples of (un)stratified comprehension

Stratified comprehension lets us form things like emptyset, universal set, set of nonempty sets, or set of subsets:

$$\{a_i \mid \bot\} : i{+}1 \quad \{a_i \mid \top\} : i{+}1 \quad \{a_i \mid \exists b_{i\text{-}1}.b_{i\text{-}1} \in a_i\} : i{+}1$$

$$\{a_i \mid a_i \subseteq a'_i\} : i{+}1 \;\; \text{where} \;\; a_i \subseteq a'_i \stackrel{def}{=} \forall b_{i\text{-}1}.(b_{i\text{-}1}{\in}a_i \Rightarrow b_{i\text{-}1}{\in}a'_i)$$

Above, we indicate levels with subscripts.

Stratified comprehension blocks $R = \{a \mid a \notin a\}$ because we can never make $i = i{+}1$.

$$\{a_i \mid \neg(a_i \in a_i)\} \quad \longleftarrow \textit{unstratified!}$$

Examples of (un)stratified comprehension

Note that TST+ sets are HOL-set-flavoured, not ZF-set-flavoured, where level i corresponds to $\overbrace{(\iota \to o) \to \cdots \to o}^{i \text{ times}}$.

You can't form a set of subsets-or-elements-of like this

$$\{a_i \mid a_i \subseteq a'_i \vee a_i \in a'_i\} \quad \longleftarrow \textit{unstratified}!$$

Put another way: the TST+ sets hierarchy is *iterative*, not *cumulative*.

Discussion of TST+ axioms

- Extensionality says sets with equal elements are equal sets.
- Comprehension says any set you can *describe* by a stratified predicate, exists.
- Typical Ambiguity is a *some/any* symmetry property: a closed ϕ valid at *some* level, is valid at *all* levels. (If I were naming the property now, I might call it level-symmetry or -invariance for closed predicates.)

In a nutshell:

- Typed set theory (TST) =
  FOL + extensional $\in$ + stratified comprehension.
- TST+ = TST + TA.

It's easy to build a sets model of TST (coming in two slides' time) but first:

Why care about ConNF?

- NF is minimal and thus in some sense canonical. Arguably, NF is what naive sets is trying to be.
- ConNF or ¬ConNF would locate more precisely the "inconsistency boundary" between naive sets and a more heavily-typed system like HOL.
- NF permits a universal set $\{a \mid \top\}$.
  We can talk about "a set of all sets" (type-theorists think: Type : Type). Freedom from hierarchies of (type) universes!
- It tells us it's OK to just have sets (and nothing but): NFU, a relative of NF that admits *urelemente* (non-set elements), is consistent. This sacrifices the idea that "everything is a set". NF is faithful to the original intuition of "everything is a set", and ConNF can be read as saying "and that's OK".

$\mathcal{V}$: the full sets hierarchy model of TST

Define the (full) sets hierarchy $\mathcal{V} = (V_0, V_1, \ldots)$ by:

$$V_0 = \mathbb{N} \quad V_{i+1} = \mathcal{P}(V_i) \quad \text{so} \quad V_i = \mathcal{P}^i(\mathbb{N})$$

So $x \in V_{i+1}$ just when $x \subseteq V_i$.

Interpret a_i to range over elements of V_i, and interpret $b_{i\text{-}1} \in a_i$ to mean *"the denotation of b is an element of the denotation of a"*.

(If you've used dependent types then this may remind you of type universes Type_0, Type_1, ... It's much the same thing.)

Problem: $\mathcal{V}$ has extensionality and comprehension, but not TA: it's not necessarily the case that $\phi \Leftrightarrow \phi^+$ (e.g. "The universe is countable" holds for V_0, but not for V_1).

Yet absence of a model of TST+ is not proof of absence. We've been stuck on this since 1937.

My claimed proof: preliminaries

- TST+ syntax is many-sorted FOL with sorts/types $\mathbb{N} = \{0, 1, 2, \dots\}$ and stratified $\in$.
- The term language at each sort i is just variables $a, b, c, \dots$.
- Thus, we have $\bot, \top, \neg, \wedge, \vee, \forall, \exists, \in$ and we only form $b{\in}a$ when $lev(a) = lev(b){+}1$.
- Write $\sim\phi$ for the (standard) *de Morgan dual* of ϕ. For example:

$$\sim\bot = \top \quad \sim(\phi\wedge\phi') = (\sim\phi)\vee(\sim\phi') \quad \sim\exists a.\phi = \forall a.\sim\phi$$
$$\sim\neg\phi = \phi \quad \sim(b{\in}a) = \neg(b{\in}a)$$

My claimed proof: preliminaries

- For ϕ closed, write ϕ^{+n} for a copy of ϕ obtained by raising the levels of all its variable symbols by n.
- For ϕ closed, write $\models^{\Box} \phi$ when $[\![\phi^{+n}]\!]$ holds in the full sets hierarchy model $\mathcal{V}$, for every n.
  E.g.: $\models^{\Box} \forall b.\exists a.b{\in}a$ (take $[\![a]\!] = \{[\![b]\!]\}$).
- Note that $\models^{\Box}$ holds for (predicates representing) comprehension, extensionality, and 'there exist at least i distinct elements' for any finite i.

Derivation system

$$\frac{}{F, \bot \vdash}\ (\bot\mathbf{L}) \qquad \frac{}{F, b{\in}a, \neg(b{\in}a) \vdash}\ (\mathbf{Ax})$$

$$\frac{F, \phi[a{:=}a'] \vdash}{F, \forall a.\phi \vdash}\ (\forall\mathbf{L}) \qquad \frac{F, \phi \vdash \quad (a \text{ fresh for } F)}{F, \exists a.\phi \vdash}\ (\exists\mathbf{L})$$

$$\frac{F, \phi, \phi' \vdash}{F, \phi\wedge\phi' \vdash}\ (\wedge\mathbf{L}) \qquad \frac{F, \phi \vdash \quad F, \phi' \vdash}{F, \phi\vee\phi' \vdash}\ (\vee\mathbf{L})$$

$$\frac{F, \sim\phi \vdash}{F, \neg\phi \vdash}\ (\neg\mathbf{L})$$

$$\frac{F, \phi \vdash \quad (\phi \text{ closed}, \ \vDash^{\Box} \phi)}{F \vdash}\ (\Box) \qquad \frac{F, \phi^{+} \vdash \quad (\phi \text{ closed})}{F, \ \phi \vdash}\ (\mathbf{Shift})$$

Derivation system: FOL + ($\Box$) + (**Shift**)

Read $F \vdash$ as 'F entails $\bot$'. It's just a FOL with empty right sequents, which is a small trick to reduce cases in a subsequent cut-admissibility argument.

($\Box$) and (**Shift**) are new:

$$\frac{F, \phi \vdash \quad (\phi \text{ closed}, \ \vDash^{\Box} \phi)}{F \vdash} (\Box) \qquad \frac{F, \phi^{+} \vdash \quad (\phi \text{ closed})}{F,\ \phi \vdash} (\textbf{Shift})$$

($\Box$) is an axiom rule, introducing any predicate valid throughout $\mathcal{V}$ (including extensionality & comprehension). As written it's undecidable — no problem for a consistency proof, but if we want to compute derivations we could probably restrict it to just extensionality, comprehension, and 'the universe has at least n distinct elements' for every $n \in \mathbb{N}$.

(**Shift**) gives us Typical Ambiguity: if closed ϕ is in the context, we can introduce ϕ^{+}.

Derivation system

Theorem 1: $\vdash$ is consistent: $\neg(\varnothing \vdash)$.

Proof: We check that every rule is sound, as follows:

Soundness

If F is a collection of predicates, write $\mathrm{Orb}(F)$ for the least collection of predicates that contains F and is such that $\phi \in \mathrm{Orb}(F)$ if and only if $\phi^+ \in \mathrm{Orb}(F)$.

In words: $\mathrm{Orb}(F)$ is the closure of F under the action of TA.

Soundness states that for each of the derivation-rules above — schematically

$$\frac{F_1 \quad \dots \quad F_n}{F}$$

— then

- if $\exists$ valuation ς to $\mathcal{V}$ such that $[\![\mathrm{Orb}(F)]\!]_\varsigma$ holds in $\mathcal{V}$,
- then $\exists\ 1 \leq i \leq n$ and valuation ς_i such that $[\![\mathrm{Orb}(F_i)]\!]_{\varsigma_i}$ holds.

In words: if *everything* below the line is possible (modulo TA), then *something* above the line is possible (modulo TA).

Are we done? Is that it? No!

Rule ($\Box$) gives us extensionality, comprehension, and typical ambiguity. Soundness gives us consistency. Fab! Are we done?

No yet; this is not enough.

To build a model and prove ConNF we need a consistent set $\mathcal{Q}$ that is, in addition to the above, *maximal* and *witnesses disjuncts and existentials*:

- $\phi \vee \phi' \in \mathcal{Q}$ must imply $\phi \in \mathcal{Q}$ or $\phi' \in \mathcal{Q}$.
- $\exists a.\phi \in \mathcal{Q}$ must imply $\phi[a{:=}a'] \in \mathcal{Q}$ for some a'.

Thus $\phi \vee \phi'$ really does mean 'ϕ or ϕ'', and similarly for $\exists a.\phi$.

Obtaining this is based on two further tricks. Call them 'Trick 1' and 'Trick 2':

Trick 1: the shift-offset Cut rule

$$\frac{F, \phi \vdash \quad G, {\sim}\phi^{+n} \vdash \quad (fv(\phi) = \varnothing \vee n = 0)}{F, G \vdash} \;(\mathbf{Cut})$$

If we could prove shift-offset Cut above is an admissible rule, then we'd be done.

Why? Because if (**Cut**) is admissible then $F, \phi \vdash$ and $F, {\sim}\phi \vdash$ implies $F \vdash$ and by the contrapositive, $F \nvdash$ implies $F, \phi \nvdash$ or $F, {\sim}\phi \nvdash$.

This enables us to saturate a finite consistent set to a maximal consistent set that witnesses disjunctions and existentials, by enumerating ϕ and adding either ϕ or ${\sim}\phi$.

E.g. if $F \nvdash$ and then $F, \phi \vee \phi' \nvdash$ then by ($\vee$**L**) also $F, \phi \vee \phi', \phi \nvdash$ or $F, \phi \vee \phi', \phi' \nvdash$, and we can extend F accordingly.

Trick 2: *partial* Cut-admissibility

Shift-offset Cut is not admissible in general:

$$\frac{F, \phi \vdash \quad G, {\sim}\phi^{+n} \vdash \quad (fv(\phi) = \varnothing \vee n = 0)}{F, G \vdash}\ (\mathbf{Cut})$$

However, partial admissibility will suffice:

Theorem 2: Shift-offset Cut is admissible in two special cases:

1. If $n = 0$. (So shift-offset Cut $\rightarrow$ normal Cut.)
2. If $n \neq 0$ and $F \cup G \cup \{\phi\}$ contains only closed predicates.

Proof: See `https://arxiv.org/pdf/1406.4060v8.pdf`, in particular Subsection 6.5 and page 26.

Admissibility of shift-offset Cut

Note that (**Shift**) / Typical Ambiguity only act on closed predicates.

Our model reflects this by consisting of a typically ambiguous *closed spine*, and an *open body* that is not. The special cases of cut-admissibility correspond to treating these two aspects of the model, separately.

My previous attempts to prove ConNF tried to directly build models that may have been too symmetric: in some sense I was trying to have shift everywhere, prove cut-admissibility everywhere, such that each level was *fully symmetric* with the level above.

This new method, which permits asymmetries during the construction, seems to be easier to work with.

Conclusions

This may be a proof of consistency of NF.

I welcome review and discussion, and proposals to formalise the argument in a theorem-prover.

$\Omega = \{a \mid \top\}$ cheers for having a universal type!