New Foundations is consistent

Sky Wilshaw

July 2023

Abstract

We give a self-contained account of a version of Holmes' proof [2] that Quine's set theory *New Foundations* [4] is consistent. This is a 'deformalisation' of the formal proof written in Lean at [7].

Note: At the present time, the formal proof [7] is incomplete, and this paper reflects the unfinished state of that proof. We aim to keep this paper proof in line with the formal proof, although as the project is ongoing, some variance is to be expected.

Contents

1 The theories at issue

In 1937, Quine introduced *New Foundations* (NF) [4], a set theory with a very small collection of axioms. To give a proper exposition of the theory that we intend to prove consistent, we will first make a digression to introduce the related theory TST, as explained by Holmes in [2]. We will then describe the theory TTT, which we will use to prove our theorem.

1.1 The simple theory of types

The *simple theory of types* (known as *théorie simple des types* or TST) is a first order set theory with several sorts, indexed by the nonnegative integers. Each sort, called a *type*, is comprised of *sets* of that type; each variable x has a nonnegative integer $\mathtt{type}(`x')$ which denotes the type it belongs to. For convenience, we may write x^n to denote a variable x with type n.

The primitive predicates of this theory are equality and membership. An equality '$x = y$' is a well-formed formula precisely when $\mathtt{type}(`x') = \mathtt{type}(`y')$, and similarly a membership formula '$x \in y$' is well-formed precisely when $\mathtt{type}(`x') + 1 = \mathtt{type}(`y')$.

The axioms of this theory are extensionality

$$\forall x^{n+1}.\, \forall y^{n+1}.\, (\forall z^n.\, z^n \in x^{n+1} \leftrightarrow z^n \in y^{n+1}) \to x^{n+1} = y^{n+1}$$

and comprehension

$$\exists x^{n+1}. \forall y^n. (y^n \in x^{n+1} \leftrightarrow \varphi(y^n))$$

where φ is any well-formed formula, possibly with parameters.

Remarks. (i) These are both axiom schemes, quantifying over all type levels n, and (in the latter case) over all well-formed formulae φ.

(ii) The inhabitants of type 0, called *individuals*, cannot be examined using these axioms.

(iii) By comprehension, there is a set at each type that contains all sets of the previous type. Russell-style paradoxes are avoided as formulae of the form $x^n \in x^n$ are ill-formed.

1.2 New Foundations

New Foundations is a one-sorted first-order theory based on TST. Its primitive propositions are equality and membership. There are no well-formedness constraints on these primitive propositions.

Its axioms are precisely the axioms of TST with all type annotations erased. That is, it has an axiom of extensionality

$$\forall x. \forall y. (\forall z. z \in x \leftrightarrow z \in y) \to x = y$$

and an axiom scheme of comprehension

$$\exists x. \forall y. (y \in x \leftrightarrow \varphi(y))$$

the latter of which is defined for those formulae φ that can be obtained by erasing the type annotations of a well-formed formula of TST. Such formulae are called *stratified*. To avoid the explicit dependence on TST, we can equivalently characterise the stratified formulae as follows. A formula φ is said to be stratified when there is a function σ from the set of variables to the nonnegative integers, in such a way that for each subformula '$x = y$' of φ we have $\sigma(`x') = \sigma(`y')$, and for each subformula '$x \in y$' we have $\sigma(`x') + 1 = \sigma(`y')$.

Remarks. (i) It is important to emphasise that while the axioms come from a many-sorted theory, NF is not one; it well-formed to ask if any set is a member of, or equal to, any other.

(ii) Russell's paradox is avoided because the set $\{x \mid x \notin x\}$ cannot be formed; indeed, $x \notin x$ is an unstratified formula. Note, however, that the set $\{x \mid x = x\}$ is well-formed, and so we have a universe set.

(iii) The infinite set of stratified comprehension axioms can be described with a finite set; this is a result of Hailperin [1].

(iv) Specker showed in [5] that NF disproves the Axiom of Choice.

While our main result is that New Foundations is consistent, we attack the problem by means of an indirection through a third theory.

1.3 Tangled type theory

Introduced by Holmes in [3], *tangled type theory* (TTT) is a multi-sorted first order theory based on TST. This theory is parametrised by a limit ordinal λ, the elements of which will index the sorts. As in TST, each variable x has a type that it belongs to, denoted `type`('x'). However, in TTT, this is not a positive integer, but an element of λ.

The primitive predicates of this theory are equality and membership. An equality '$x = y$' is a well-formed formula when $\mathtt{type}(`x') = \mathtt{type}(`y')$. A membership formula '$x \in y$' is well-formed when $\mathtt{type}(`x') < \mathtt{type}(`y')$.

The axioms of TTT are obtained by taking the axioms of TST and replacing all type indices in a consistent way with elements of λ. More precisely, for any order-embedding $s : \omega \to \lambda$, we can convert a well-formed formula φ of TST into a well-formed formula φ^s of TTT by replacing a type variable α with $s(\alpha)$.

Remarks. (i) Membership across types in TTT behaves in some quite bizarre ways. Let $\alpha \in \lambda$, and let x be a set of type α. For any $\beta < \alpha$, the extensionality axiom implies that x is uniquely determined by its type-β elements. However, it is simultaneously determined by its type-γ elements for any $\gamma < \alpha$. In this way, one extension of a set controls all of the other extensions.

(ii) The comprehension axiom allows a set to be built which has a specified extension in a single type. The elements not of this type may be considered 'controlled junk'.

We now present the following striking theorem.

Theorem 1.1 (Holmes)**.** NF is consistent if and only if TTT is consistent.

The proof is not long, but is outside the scope of this paper; it requires more model theory than the rest of this paper expects a reader to be familiar with, and relies on additional results such as those proven by Specker in [6].

Thus, our task of proving NF consistent is reduced to the task of proving TTT consistent. We will do this by exhibiting an explicit model (albeit one that requires a great deal of Choice to construct). As TTT has types indexed by a limit ordinal, and sets can only contain sets of lower type, we can construct a model by recursion over λ. This was not an option with NF directly, as the universe set $\{x \mid x = x\}$ would necessarily be constructed before many of its elements.

2 Outline

References

[1] Theodore Hailperin. "A set of axioms for logic". In: *Journal of Symbolic Logic* 9.1 (1944), pp. 1–19. DOI: 10.2307/2267307.

[2] M. Randall Holmes. *NF is Consistent*. 2023. arXiv: 1503.01406 [math.LO].

[3] M. Randall Holmes. "The Equivalence of NF-Style Set Theories with "Tangled" Theories; The Construction of ω-Models of Predicative NF (and more)". In: *The Journal of Symbolic Logic* 60.1 (1995), pp. 178–190. ISSN: 00224812. URL: http://www.jstor.org/stable/2275515.

[4] W. V. Quine. "New Foundations for Mathematical Logic". In: *American Mathematical Monthly* 44 (1937), pp. 70–80. URL: https://api.semanticscholar.org/CorpusID:123927264.

[5] Ernst P. Specker. "The Axiom of Choice in Quine's New Foundations for Mathematical Logic". In: *Proceedings of the National Academy of Sciences of the United States of America* 39.9 (1953), pp. 972–975. ISSN: 00278424. URL: http://www.jstor.org/stable/88561.

[6] Ernst P. Specker. "Typical Ambiguity". In: *Logic, Methodology and Philosophy of Science*. Ed. by Ernst Nagel. Stanford University Press, 1962, pp. 116–123.

[7] Sky Wilshaw et al. *The consistency of New Foundations*. 2022–2023. URL: https://leanprover-community.github.io/con-nf/.