Realizability and the Possibility of a Consistency Proof for the Constructive Fragment of Quine's Set Theory NF

Thomas Forster

Queens' College
and
Department of Pure Mathematics and Mathematical Statistics
Cambridge University
(Prepared for the *Facets of Realizability* Meeting in Paris, July 2019)
(slightly) revised

December 20, 2021

ABSTRACT

It is known from the work of Specker [3] that Quine's NF is consistent iff the theory T$\mathbb{Z}$T of Typed Set Theory with types indexed by $\mathbb{Z}$ remains consistent when we add the scheme of biconditionals $\phi \longleftrightarrow \phi^+$, where ϕ^+ is the result of raising all type indices in ϕ by 1. Since evidently T$\mathbb{Z}$T $\models \phi$ iff T$\mathbb{Z}$T $\models \phi^+$ it looks as if there should be realizers for the corresponding biconditionals $\phi \longleftrightarrow \phi^*$ and thereby a proof of consistency for *i*NF (the constructive fragment of NF) that is not at the same time a reason to believe in the consistency of the full classical theory. There seems to be a connection here with Visser's Logic BPC in [4].

Contents

This is very much work-in-progress, and has benefitted from discussons with Randall Holmes and Allen Hazen. It is possible that the connections with [4] will not get explored before the deadline.

This document is only very slightly altered from the abstract that was submitted for (and accepted for) the Paris *Facets of Realizability* Meeting in Paris, July 2019 which i was unable to attend beco's I was in Cambridge recovering from surgery.

1 Five grades of Typical Ambiguity

The most general setting for ideas like these is a first-order language $\mathcal{L}$ with a bijection $\sigma : \mathcal{L} \longleftrightarrow \mathcal{L}$ that commutes with quantifiers and connectives, and an $\mathcal{L}$-theory T such that σ is an automorphism of T in the sense that, for all $\phi \in \mathcal{L}$, $T \vdash \phi$ iff $T \vdash \sigma(\phi)$. (Henceforth we will write 'ψ^σ' rather than '$\sigma(\psi)$'). The two Specker articles are essential reading. T$\mathbb{Z}$T (which is the strongly typed set theory with levels indexed by $\mathbb{Z}$, the theory in [5][1]) is such a theory. Indeed T$\mathbb{Z}$T is likely to be our main preoccupation in what follows, and the automorphism of $\mathcal{L}$(T$\mathbb{Z}$T) that is of interest to us is the operation that bumps up the type of a formula by one. Traditionally we write this automorphism with a '+' symbol: thus ϕ^+ is the result of lifting all type subscripts in ϕ by one. The phrase 'typical anbiguity ' is often used in contexts like this, and it comes in five grades.

Grade (i) $T \vdash \phi$ iff $T \vdash \phi^+$;
Grade (ii) $T \cup \{\phi \longleftrightarrow \phi^+\}$ is consistent for each ϕ;
Grade (iii) $T \cup \sum_{\phi \in \mathcal{L}} \phi \longleftrightarrow \phi^+$ is consistent;
Grade (iv) $T \vdash \sum_{\phi \in \mathcal{L}} \phi \longleftrightarrow \phi^+$;
Grade (v) T has an ambiguous model.

Grade (v) is complicated to state in general but in the case of interest here (T$\mathbb{Z}$T) an ambiguous model (*glissant* in French) is a model with an automorphism that sends elements of type (level) i to elements of level $i + 1$. Any such model will give a model of NF as per Specker [3].

(It may be worth noting *en passant* that—as Specker points out—if the + operation is an involution then (ii) implies (iii). However the + operation in play in T$\mathbb{Z}$T is of infinite order so this observation is not very useful to us, tho' it is worth keeping in mind for later use in a more general context.)

Classically there is a theorem of Specker's [2] that says that grade (iii) implies grade (v). (A theory of grade (iii) can be extended to a theory of grade (iv) and grade (iv) implies grade (v) by general model-theoretic nonsense, and we supply no proof). This gives a reduction of Con(NF) to the assertion that T$\mathbb{Z}$T is grade (iii)). Altho' Marcel Crabbé showed in [1] that T$\mathbb{Z}$T is grade (ii), unfortunately there is no obvious reason to believe that it is grade (iii) (and it manifestly isn't grade (iv)). Classically you can have theories that are grade (ii) but have no extensions that are grade (iv). Specker [2] supplies examples and we will consider them below.

1.0.1 Specker's Example of a Theory that is grade (ii) but not grade (v)

From: Specker [2].

[1] In [5] this theory is called the *Theory of Negative Types* and for many years was called 'TNT'. Nowadays the notation 'T$\mathbb{Z}$T' is preferred, leaving 'TNT' to denote the analogous theory with types indexed by the negative integers. One wants to distinguish the two because it is far from clear that every model of TNT can be "extended upwards" to a model of T$\mathbb{Z}$T.

The language has types indexed by $\mathbb{Z}$ (so each variable is restricted to range over one level only), with equality but no nonlogical vocabulary. Our theory T will have two axiom schemes:

(1) There are precisely 1, 2, or 3 elements of each type;
(2) There are not equally many elements of type k and of type $k+1$.

To be formal about it let us write '$\exists! x_i$' for '$(\exists x_i)(\forall y_i)(x_i = y_i)$',
'$\exists_2! x_i$' for '$(\exists x_i, y_i)(x_i \neq y_i \wedge (\forall z_i)(x_i = z_i \vee z_i = y_i))$' and
'$\exists_3 x_i$' similarly.

At each level i, T has axioms

$$(\forall a_i\, b_i\, c_i\, d_i)(a_i = b_i \ \vee\ a_i = c_i \ \vee\ a_i = d_i \ \vee\ b_i = c_i \ \vee\ b_i = d_i \ \vee\ c_i = d_i)$$
$$\neg(\exists! x_i \wedge \exists! x_{i+1})$$
$$\neg(\exists_2! x_i \wedge \exists_2! x_{i+1})$$
$$\neg(\exists_3! x_i \wedge \exists_3! x_{i+1})$$
$$(\exists x_i)(x_i = x_i)$$

This T is a theory in (many-sorted) first-order logic, but many of its features can be captured in a propositional theory T_{prop} in a language $\mathcal{L}$ with propositional letters p_i, q_i and r_i for all $i \in \mathbb{Z}$. The theory T_{prop} has two schemes:

$$(p_i \wedge \neg q_i \wedge \neg r_i) \ \vee\ (\neg p_i \wedge q_i \wedge \neg r_i) \ \vee\ (\neg p_i \wedge \neg q_i \wedge r_i)$$
$$\neg(p_i \wedge p_{i+1}) \ \wedge\ \neg(q_i \wedge q_{i+1}) \ \wedge\ \neg(r_i \wedge r_{i+1}).$$

one instance of each for each $i \in \mathbb{Z}$. (Secretly p_i says there there is precisely one object at level i; q_i says there are precisely two ...)

Recall that we write ψ^+ for the result of increasing every subscript in ψ by 1. If f is a valuation defined on the letters in $\mathcal{L}$ then f^+ is the valuation $v^+ \mapsto f(v)$. This ensures that f satisfies ψ iff f^+ satisfies ψ^+.

T_{prop} considered as the deductive closure of these axioms (a set of formulæ), has lots of automorphisms (one can permute the letters $\{p, q, r\}$) but the automorphism of interest to us is that which sends every p-variable 'p_i' to 'p_{i+1}' and q- and r-variables similarly. We shall write this automorphism with a '+' sign. Altho', for all ψ, we have: $T_{\text{prop}} \vdash \psi$ iff $T \vdash \psi^+$, nevertheless we do not have $T_{\text{prop}} \vdash \psi \longleftrightarrow \psi^+$ for all ψ. Thus T is grade (i) but is not grade (iv).

Remark 1
(a) $T_{\text{prop}} \cup$ the scheme of biconditionals $\psi \longleftrightarrow \psi^+$ is inconsistent;
(b) Each biconditional $\psi \longleftrightarrow \psi^+$ is individually consistent with T_{prop}.

Proof:

(a) is pretty obvious; for (b) fix an arbitrary $\mathcal{L}$-formula ψ; we will find a valuation satisfying $\psi \longleftrightarrow \psi^+$. Suppose (with a view to obtaining a contradiction) that every valuation satisfies precisely one of ψ and ψ^+. Think of the valuation f that goes $\ldots p\, q\, r\, p\, q\, r \ldots$ (with period 3) as you ascend through the levels and the two valuations f^+ and f^{++}. Recall that f sat ψ iff f^+ sat ψ^+ and so on. Do any of these valuations actually satisfy $\psi \longleftrightarrow \psi^+$? If they do, we are happy. If not, then each of them satisfies precisely one of $\{\psi, \psi^+\}$. Without loss of generality f satisfies ψ but not ψ^+; then f^+ satisfies ψ^+ but not ψ, f^{++} satisfies ψ but not ψ^+ and f^{+++} satisfies ψ^+ but not ψ. But $f^{+++} = f$.

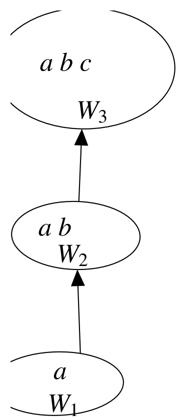
■

We observe without proof that (a) and (b) of the remark hold also for (our formulation of) Specker's original first-order theory T.

Thus, manifestly, classically grade (i) does not imply grade (iv). However, if we are thinking constructively, then the + operation on proofs provides obvious candidates for realizers of the biconditionals in (ii) and (iii) and thereby gives us reason to believe that *constructively* grade (i) implies grade (iv); thus we are looking for a version of Specker's theorem for constructive logic.

In propositional logic any contradiction provable classically is provable constructively, so there is no hope for a version of Specker's theorem for constructive *propositional* logic. However, there may be a version for constructive *first-order* logic. For example: it may be that the first-order theory T (which, remember, was Specker's original example) has a constructive version to which we can consistently add an axiom scheme of typical ambiguity. (In fact I cunningly chose the axiomatisation above with precisely this possibility in mind, so that no changes are needed.)

It turns out that this is indeed the case: if we use a constructive logic then this T remains consistent when we add the scheme $\phi \longleftrightarrow \phi^+$. If we then discard the type subscripts on the variables we obtain a one-sorted theory that says there is at least one but no more than three distinct things, there is not precisely one thing, nor precisely two nor precisely three. This theory is clearly satisfied by the possible world model with three worlds W_1 (the root world) which contains a and can see (itself and) W_2 ... which contains a and b, whch in turn can see (itself and) W_3 which contains a, b and c, and can see only itself.



The above possible world model is the grade (v) object we are looking for.

This inspires hope in the following possibility:

CONJECTURE 1
Whenever T is a constructive theory in a first-order language admitting an automorphism σ of T such that
(i) $T \vdash \psi$ iff $T \vdash \sigma(\psi)$ and
(ii) the scheme $\psi \longleftrightarrow \psi^{\sigma}$ is consistent relative to T
then T has a model admitting an automorphism corresponding to σ.

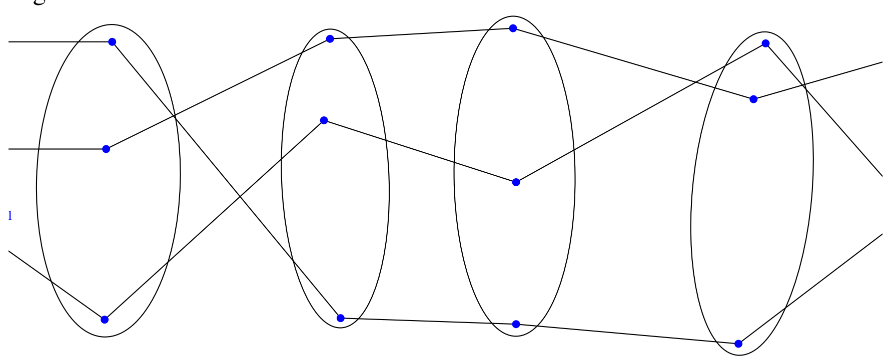
The reason for interest in this situation is that there might be classical theories which do not satisfy the antecedent of Specker's theorem (alluded to above) that says that grade (iii) implies grade (v) but whose constructive fragments satisfy the assumptions of the conjecture. Specifically the thought is that T$\mathbb{Z}$T might be such a theory—and that this might matter. And it might matter beco's it may yet be that Holmes' consistency proof for NF falls at the final fence, and in such an eventuality it would be nice if nevertheless the constructive fragment admitted a consistency proof along the above lines.

To clarify the situation and estimate the hopes for such a result we need to think a bit about what realizers are and what it is to raise the type of a formula.

2 Type-raising and Realizers

The thought that launched this discussion was the idea that realizers for formulæ in a strongly typed language such as $\mathcal{L}$(T$\mathbb{Z}$T) were the kind of thing that type-raising operation could act on. If they are, then the raising of types certainly provides realizers for conditionals like $\phi \to \phi^+$.

We need to start by reflecting that there is no good notion of type-raising in propositional languages. And this is despite the artful way in which the reader was tricked into accepting T_{prop} as a typed propositional theory. Consider the (typical) axiom $\neg(p_i \wedge p_{i+1})$. The trickery is the exploitation of the subscripts. As far as the propositional language is concerned the propositional letters have no internal structure: every permutation of the set of propositional letters of a propositional language is an automorphism of the language. One would expect that any permutation of the propositional letters would fix T_{prop} at least in the sense of sending it to something α-equivalent to it, but this is not so. If one thinks of T_{prop} as expressed in a language where none of the propositional letters have internal structure then T_{prop} in effect defines some structure on that language. It defines a three-place relation (the ellipses in the picture below) and a two-place relation (the left-to-right edges in the picture below). For a permutation of the set of propositional letters to fix T_{prop} (at least in the sense of sending it to something α-equivalent to it) then it must preserve the ellipses and the injective character of the edges.



But we don't need to make sense of type-raising in propositional logic, for reasons noted above. However we do need to think about precisely what we mean by "raising the types" in a *first-order* formula. Now a first-order theory is a set of *closed* formulæ so—*prima facie*—we only need to consider what it is to "raise the type" of a closed formula. Thus this + operation, if taken in the way explained above, is really an operation defined not on (closed) formulæ but on α-equivalence classes of (closed) formulæ. That is just as well, because the idea of raising types in an open formula is problematic. The simplest open formula is of course a naked variable. What happens, what does one get, if one raises the type of a variable? How do we decide which variable of the next level to replace a given variable with? If we are raising the type of a variable as part of an act of raising the type of a closed formula then it doesn't matter; the axioms are—all of them—closed formulæ, and all the results of performing this (somewhat nondeterministic) operation on ϕ will be α-equivalent to each other. If we are trying to raise the type of a naked variable it matters a great deal.

bad break

In a many-sorted language each variable has a sort, and one has to have a way of ascertaining what that sort is. However it is not part of the spec of a many-sorted language that there should be a bijection between the set of variables of one sort and the set of variables of another. Consider the theory of Vector Spaces, sometimes set up as a two-sorted theory, with Greek letters for variables to range over scalars, and Roman letters for variables to range over vectors. The language is not equipped with any bijection between the vector and the scalar variables.

H I A T U S

If type-raising is to be a realizer for $\phi \to \phi^+$ then realizers for ϕ have to be fairly syntactic objects - things that + can operate on. But + doesn't work very well on things with free variables. So it looks as tho' things with free variables can't have realizers. Or perhaps (and this would be a good outcome) things with free variables can have realizers all right, it's just that type-raising on them doesn't work.

H I A T U S

The usual way of thinking of variables in the context of typed set theory (TST, T$\mathbb{Z}$T...) is to take a variable to be a letter from some standard alphabet (typically the Roman alphabet) decorated with a subscript (or, in the early literature, a superscript) from $\mathbb{N}$ or $\mathbb{Z}$. If we think of variables that way then we can define the result of raising the type of a variable[2] v as the result of adding 1 to the numerical subscript in v. Then we define ϕ^+ to be the result of raising all the subscripts on all the variables in ϕ by precisely 1, with the result that + is an operation defined on formulæ themselves rather than merely on their α-equivalence classes.

Now for the second question: is it a good idea? The answer to this is: possibly. For one thing, this is the way it has always been done, the people who set this stuff up were not mugs—they knew what they were doing—and no-one has ever complained. However, not everything in the garden is rosy: there is an infelicity and an invitation to error.

Well, *two* infelicities, actually.

The first infelicity is that the presence of sort subscipts clogs up a place which has a long and honourable history of serving other useful purposes, such as notating sequences. This is notationally annoying but not mathematically substantial. We will not consider it further.

The second infelicity is of a different nature altogether. The attachment of subscripts provides extra information about the variables which—from the point of $\mathcal{L}$(T$\mathbb{Z}$T)—is spurious: the connection between 'x_2' and 'x_3' that is visible to us is not visible to the language. As Randall Holmes points out[3] '$\{x_0 : x_0 \notin x_1\}$' is a well-formed formula (at least once one had added set abstraction) but it does not capture the Russell class at level 1; the '$\{x_0$:' that binds the variable of level 0 cannot bind the variable of level 1. This kind of internal structure of the variables has—and can have—no semantics [4]. It is this second infelicity that offers us the invitation to error. According to the new

[2]Here 'v' is of course a variable ranging over variables(!)

[3]in conversation.

[4]There is a parallel here with a question on a venerable example sheet that i have to teach: "supply a

definition of the operation + it is defined on formulæ with free variables. On the face of it that ought to be harmless, so how is it an invitation to error?

As remarked above any T$\mathbb{Z}$T-proof of ϕ can be turned into a proof of ϕ^+. This means that "from $\vdash \phi$ infer $\vdash \phi^+$" is an admissible[5] rule of T$\mathbb{Z}$T. *However this is legitimate only if ϕ is a closed formula.* and there are no live assumptions in the proof. If + is defined only on closed formulæ then the admissible rule is applicable to closed formulæ only. Once we extend + to formulæ with free variables then new formulæ come within the purview of the rule and it may cease to be admissible. This is in fact what happens.

rewrite below here

[outline: if we are allowed the rule on formulæ with free variables on ce end up proving the ambiguity scheme]

If one thinks that it applies to formulæ containing free variables, then one might feel one can prove that any given level has sets of all finite sizes and is therefore infinite, and one would do it roughly as follows. If level l contains n distinct things, then level $l-1$ also contains n distinct things, whereupon level l will suddenly be found to contain 2^n distinct things. Thus we can power an induction on n to prove in the theory that, for all n, level i has at least n elements. In fact this purported proof needs an extra feature beyond applying the ambiguity rule to formulæ with free variables. Consider the two formulæ

$$\text{Level } i \text{ contains at least } n \text{ distinct things} \qquad (i)$$

and

$$\text{Level } i+1 \text{ contains at least } n \text{ distinct things} \qquad (i+1)$$

We can arrange for formula $i+1$ to be + of formula (i)—so that sentence $i+1$ implies sentence i—if n is a concrete numeral, something of the kind that can be captured with numerically definite quantifiers; but then the 'n' is not a variable, and cannot be something on which we are doing an induction. On the other hand if 'n' is a variable then what is its type? And where is the induction going on? In T$\mathbb{Z}$T? Or in the metalanguage?

The idea is that this induction establishes that the natural numbers of level i is well-behaved in the sense of having no last element, so we have proved the axiom of

formula that is true in all structures with at least n elements". The students want to write

$$(\exists x_1 \ldots x_n)(\bigwedge_{1 \leq i \neq j \leq n} x_i \neq x_j)$$

and of course in some sense they are right to want to write it. However the quoted formula is not a wff of first-order logic, or at least not of the first-order logic they have been told about—and this point needs to be made to them. I find myself telling them that it is perhaps a program that evaluates to such a formula on being given the input n.

[5] Wikipædia puts it very succinctly:"In logic, a rule of inference is admissible in a formal system if the set of theorems of the system does not change when that rule is added to the existing rules of the system. In other words, every formula that can be derived using that rule is already derivable without that rule, so, in a sense, it is redundant. The concept of an admissible rule was introduced by Paul Lorenzen (1955)."

infinity. Of course one cannot in fact do anything of the sort (tho' one can prove in the metatheory that for every *concrete numeral* n, level i has at least n elements). From time to time incautious workers make this mistake; to the best of my knowledge no such attempt has made it into print, but it's a permanent possibility

Another version. Recall the definition of finite beth numbers: $\beth_0 = 1$; $\beth_{n+1} = 2^{\beth_n}$.

Then we want to say: if there are at least $\beth_n$ things at level i then there are at least $\beth_{n+1}$ things at level $i + 1$. (This by UG on n and i). Then, by typical ambiguity, there are at least $\beth_{n+1}$ things at level i. Therefore we have proved: "If there are at least $\beth_n$ things at level i then there are at least $\beth_{n+1}$ things at level i." So, by induction on n, for all n there are at least $\beth_{n+1}$ things at level i.

"But!" the reader may think "if every level of a model $\mathfrak{M}$ of T$\mathbb{Z}$T is infinite then every natural number (thought of as an equinumerosity class) at every level must be inhabited. And that is the axiom of infinity!" This line of reasoning certainly proves that every concrete natural number of $\mathfrak{M}$ is inhabited (is nonempty) but, for the axiom of infinity to hold, we need to show the nonemptiness of every member of the possibly larger set of natural-numbers-according-to-$\mathfrak{M}$. And something is a natural-number-according-to-$\mathfrak{M}$ if it belongs to every set of $\mathfrak{M}$ that contains 0 and is closed under successor. If $\mathfrak{M}$ doesn't contain very many sets that contain 0 and are closed under successor then the intersection of all of them might be too large, and—as a result—contain things that are not concrete natural numbers.

The discussion in the preceding paragraph doesn't show that there is no proof of the axiom of infinity in T$\mathbb{Z}$T. For that we need a compactness argument.

References

[1] Marcel Crabbé "Typical Ambiguity and the Axiom of Choice" *Journal of Symbolic Logic* **49** (1984) pp. 1074–1078

[2] Ernst Specker "Dualität" *Dialectica* **12** pp. 451–465. English translation available on `https://www.dpmms.cam.ac.uk/~tf/dualityquinevolume.pdf`

[3] Ernst Specker "Typical Ambiguity" in Logic, Methodology and Philosophy of Science *ed* Nagel *et al* Stanford 1960 pp 116–124.

[4] Albert Visser. A Propositional Logic with explicit Fixed Points, Studia Logica **40** (1981), 155–175.

[5] Wang, H. "Negative Types" *MIND* **61** pp. 366–8.