NF15104Holmes on Realizability for *i*NFsection.15

Notes on Constructive NF

Thomas Forster

August 2, 2021

Contents

Introduction

By 'constructive NF' I mean the system of set theory that has the same nonlogical axioms as NF but is embedded in an intuitionistic logic instead of a classical logic. Since in this weakened logic certain classical equivalences do not hold, no harm can be done by spelling out the axioms in detail. We have the axiom of extensionality in the form

$$(\forall xy)(x = y \longleftrightarrow (\forall z)(z \in x \longleftrightarrow z \in y))$$

and a scheme of comprehension axioms

$$(\forall \vec{x})(\exists y)(\forall z)(z \in y \longleftrightarrow \phi(\vec{x}, z))$$

where ϕ is weakly stratified and has no occurrences of 'x'.

Over the years various people have thought that constructive versions of NF might be easier to attack than the full classical theory, and although a small amount of progress has been made in our understanding of the situation, there is no adequate summary of known results available. Every now and then brief articles are submitted to journals, but none of them contain any significant results. There is a good reason for this: no significant results are known! The purpose of this document is to summarise the basic facts that the various workers (mainly Holmes, Dzierzgowski and me) have been able to ascertain, in order to ensure that what little is known is all in one readily accessible place. Most of the remarks and lemmas below are unattributed. This is not because I am claiming that they were proved first by me—the large majority of them were not—but because I cannot now remember who proved them first! As always when working on any aspect of NF involving intuitionism or proof theory, I am greatly endebted to Randall Holmes, Marcel Crabbé (as always) and Daniel Dzierzgowski. Daniel Dzierzgowski has also done a great deal of very interesting work on intuitionistic versions of the type thory that underlies NF, which I do not discuss here. I am also endebted to Jan Ekman, and to the proof theorists and intuitionists of the Computer Laboratory in Cambridge, particularly Peter Johnstone (the extent of whose rôle in my education will become clear in what follows). Others who feature in the correspondence are John Bell, Sergei Tupailo and Carsten Butz. I am endebted to Holmes also for permission to include his chapter in this tutorial as section 14.

This tutorial has—to the extent that is always inevitable in situations like these—the character of the briefing paper that its author would have liked to have at the outset. The one person for whom this paper was written no longer exists! Notwithstanding that I am grateful to Marcel Crabbé for the suggestion that I write it up for the NF 70th anniversary volume, and also for the opportunity this affords me to straighten out my thoughts. I am uncomfortably aware that the document that the reader has before them is clearly a work-in-progress (now very different from the text that appeared in the NF 70th anniversary volume) and I propose to maintain and update it, and make it available from my home page.

There is a slight problem with nomenclature. Naturally there was a debate about what the constructive version of NF should be called. All the obvious candidates for names for this system have obvious disadvantages and I will not tire the reader by recounting them. Maurice Boffa said the system should be called 'INF'. He was my *Doktorvater*, and he is now dead, so he cannot be argued with: 'INF' it is. Or was, until recently. Albert Visser suggested 'iNF' which sounded pretty good to me, but later i decided that '*i*NF' sounded even better. Beeson doesn't like it: he doesn't want it to sound like an Apple product. Why not? Apple records was the Beatles' recording company and i'm a Beatle fan; Beeson should know better.

Interest in intuitionistic versions of NF dates back to my Ph.D. thesis. Of course my primary interest there was in proving the consistency of NF itself, but I was attracted by the idea of doing some forcing semantics in the following way. If $\mathfrak{M}$ is a model of Russellian simple type theory, let $\mathfrak{M}^n$ be the result of chopping the bottom n types off $\mathfrak{M}$ and relabelling appropriately. Fix a model $\mathfrak{M}$ of Russellian simple type theory and consider the family of all structures

$$(\prod_{i \in \mathbb{N}} \mathfrak{M}^i)/F$$

where 'F' varies over the nonprincipal filters over $\mathbb{N}$. If F is ultra one obtains a model of simple type theory; the (optimistic) thought being that this construction will smear out those differences between types that were violating typical ambiguity. Naturally this was never going to give a model of NF, because Łoś's theorem would ensure that all the pathologies demonstrable in NF would have to be put into $\mathfrak{M}$ to begin with, but in my thesis I considered what one might achieve by turning the above family into a Kripke model of constructive typed set theory by equipping it with the inclusion relation (on the filters) for an accessibility relation. This gives us a Kripke model $\mathfrak{M}$ of an intuitionistic version of this simple type theory with a weak polymorphism: $\mathfrak{M} \models \phi$ iff $\mathfrak{M} \models \phi^+$, but not $\mathfrak{M} \models \phi \longleftrightarrow \phi^+$, which is part of what would be needed for a proof of $Con(i\text{NF})$.[1]

This document is a lineal descendent of ("identical by descent") and supercedes [21].

1 Background Expectations

The obvious questions to ask about *i*NF are: Is it consistent? Does the consistency of the classical theory follow easily from the consistency of the constructive theory by some sort of negative interpretation *à la* Powell [28]? What is the constructive content of Specker's proof of the axiom of infinity? Can we implement Heyting Arithmetic in *i*NF?

The obvious method for constructing for NF an analogue of the Powell negative interpretation doesn't work, since the collection of hereditarily stable sets

[1] We actually need something slightly stronger than this.

is defined by an unstratified formula and might not be a set. We need it to be a set because *i*NF believes there is a universal set!

There is a proof of the axiom of infinity in the classical version of NF. To this day no-one has ever ascertained the precise constructive content of this proof, but we have been able to obtain a significant result even without doing so. In brief, if V is Kuratowski finite then Ω is Kuratowski finite, which implies classical logic, which enables us to run Specker's proof that V is not Kuratowski finite. So we have proved in *i*NF that V is not Kfinite without discovering any constructive content to Specker's proof! (See below for definitions and details). Suppose we were to ascertain the constructive content of Specker's proof, would this help? Suppose our luck were well and truly in and the proof were entirely constructive .... Unfortunately this gets us nowhere: the mere fact that V is not Kfinite is not—in an intuitionistic context—sufficient for there to be an implementation of Heyting arithmetic in *i*NF. To implement Heyting arithmetic we need cardinals of Nfinite sets, and the mere fact that V is not Kfinite doesn't seem to imply that every Nfinite set has inhabited complement, which is what we would need to implement Heyting's arithmetic. Classically one can prove that if there is even one set that is not finite then there is a smooth implementation of arithmetic. (Apparently this fact was known to Frege and—in some circles—is even known as "Frege's theorem".) At present the situation seems to be that the constructions available in the classical case that give us implementations of Heyting arithmetic clearly fail—and for well understood reasons—and there is no obvious place to look for replacements.

Those two items were reasons to believe that *i*NF is weaker than NF. There now follows a second batch of three items (admittedly the second and third are recondite) which are listed separately because they are reasons for supposing that *i*NF is consistent that are not at the same time reasons for supposing that NF is consistent.

Specker's equiconsistency theorem relating NF to type theory with complete ambiguity has an obvious relevance for realizability approaches, since the ambiguity axioms have obvious realizations. The situation is not completely straightforward, because the version of Specker's theorem appropriate for intuitionistic models is very hairy. The place to look for enlightenment on this is Dzierzgowski's Ph.D. thesis [18].

One way forward from this is pointed to us by a suggestion of Randall Holmes: develop a *realizability* interpretation for intuitionistic type theory! That is to say, take the interpretation of a conditional $p \to q$ to be the set of functions from interpretations of p to interpretations of q. A proposition is (constructively) authorised if its interpretation contains the denotation of a closed λ term. Now clearly the interpretation of $\phi \to \phi^+$ has one obvious member which is "raise types!".

So all we need to do is to associate to intuitionistic typed set theory a λ calculus containing a term that denotes this function. Although it is far from clear how to do this, if we were to do it we would have a consistency proof for *i*NF *that did not obviously give rise to a consistency proof for NF.* This is to be expected because all the ideas that suggest that NF should be consistent are type-theoretic and are really arguments that iNF ought to be consistent.

Finally: stratified formulæ are slightly better behaved proof-theoretically than unstratified formulæ, and constructive logic is significantly better behaved proof-theoretically than nonconstructive logic. Putting these together might enable us to find—by means of cut-elimination or something like that—a cute proof-theoretical demonstration of the consistency of iNF.

To summarise, we are still taking bets on what the status of iNF will turn out to be. Holmes thinks it is strong, I think it is weak. I think it is weak because the obvious ways to interpret classical NF and Heyting arithmetic into it both fail. Holmes thinks this is the Gods laying a false trail. The book is still open.

2 Definitions

The presentation is informal, in the sense that I do not present proofs as explicit mathematical objects. I sometimes appeal to the twin rules of $\in$-introduction and $\in$-elimination:

$$\frac{\phi(x)}{x \in \{y : \phi(y)\}}(\in\text{-int}) \qquad\qquad \frac{x \in \{y : \phi(y)\}}{\phi(x)}(\in\text{-elim})$$

but there is very little explicit natural deduction.

In connection with these two $\in$-rules it might be worth noting that if ϕ is weakly stratified then $y \in \{z : \neg\neg\phi\} \longleftrightarrow \neg\neg(y \in \{z : \phi\})$. This is true because $y \in \{z : \phi\} \longleftrightarrow \phi(y)$, so $\neg\neg(y \in \{z : \phi\}) \longleftrightarrow \neg\neg\phi(y)$ but the RHS is equivalent to $y \in \{z : \neg\neg\phi\}$.

We will also allude later to a **term rule** which is a kind of generalisation of the ω-rule in arithmetic. It allows us to infer $(\forall x)(F(x))$ from the infinitely many premisses $F(t)$ for all closed terms t. 'Closed terms' in this context means of course *weakly stratified set abstract*.

While on the subject of natural deduction, we might record the following observation of Jan Ekman's, made to me in conversation:

Remark 1 *(Ekman)*
There is no normal proof of $(\forall x)(\exists y)(y \notin x)$ *even in naïve set theory.*

Proof: Assume that there is a normal proof of this formula. We know that a normal proof ends with an introduction. Using this argument three times we infer that there is a normal deduction of $\perp$ from $t \in x$, for some term t. Since $\perp$ cannot be the conclusion of an introduction this deduction has an E-main branch.

Since $t \in x$ is the only open assumption in the deduction $t \in x$ is the topmost formula in the E-main branch. Since $t \in x$ is not the endformula of the deduction, and occurs in the E-main branch $t \in x$ is major premise of an elimination inference!

This is a contradiction. $\blacksquare$[2]

ι is the singleton function, so that $\iota``x$ is $\{\{y\} : y \in x\}$. We could write '$\iota(x)$' or '$\iota`x$' for the singleton of x but we will continue to write '$\{x\}$' as usual.
$\bot$ is the truth-value `false`;
$\emptyset$ is the empty set;
0 is the number zero.

(These are all distinct things, and deserve separate notations.)

Ω the algebra of truth-values, and
$\top$ is the truth-value `true`—is its top element.

When we wish to think of Ω concretely we can take it to be $\mathcal{P}(\{\emptyset\})$;
$[[\phi]]$ is the truth-value of ϕ (when ϕ is weakly stratified) so that—when thought of concretely—$[[\phi]]$ is $\{x : x = \emptyset \wedge \phi\}$, $\bot$ is $\emptyset$ and $\top$ is $\{\emptyset\}$.
Single square brackets—as in '$[x]_R$'—are a notation for the equivalence class of x under the equivalence relation R
$\sim\sim x =_{df} \{y : \neg\neg(y \in x)\}$; (observe that this is always a set if x is)
$\mathcal{P}_\Phi(x)$ is $\{y \subseteq x : \Phi(y)\}$;
A truth-value is **dense** iff its double complement is $\{\emptyset\}$ aka the `true`
H_Φ is $\bigcap\{y : \mathcal{P}_\Phi(y) \subseteq y\}$, namely the least fixed point for $\lambda x.\mathcal{P}_\Phi(x)$.
(The greatest fixed point, $\bigcup\{y : y \subseteq \mathcal{P}_\Phi(y)\}$), doesn't have a special notation here, though no doubt it should!)
A set x is **determinate** iff $(\forall y)(y \in x \vee y \notin x)$;
A set x is **stable** iff $(\forall y)(\neg\neg(y \in x) \to y \in x)$;
A set x is **orthogonal** iff $(\forall yz \in x)(\neg\neg(y = z) \to y = z)$;
A set x is **discrete** iff $(\forall yz \in x)(y = z \vee y \neq z)$;
A set x is **inhabited** iff $(\exists y)(y \in x)$;
A set x is **nonempty** if $\neg(\forall y)(y \notin x)$.
A **transversal** of a disjoint family is a set that meets each member of the family on a singleton.

I don't think that "orthogonal" is standard usage. It is easy to verify that the relation $\{\langle x, y\rangle : \neg\neg(x = y)\}$ is an equivalence relation. An orthogonal set is one whose every intersection with an equivalence class under this relation is either empty or is a singleton.

We say "X is **closed under adjunction**" to mean

$$(\forall x \in X)(\forall y)((x \cup \{y\}) \in X)$$

Note that we do **not** require that $y \notin x$.

The **set** of Kfinite sets is the intersection of all sets containing $\emptyset$ and closed under adjunction, thus:

$$\text{Kfin} \ = \ \bigcap\{Y : \emptyset \in Y \wedge (\forall xy)(x \in Y \to x \cup \{y\} \in Y)\}$$

[2]Curiously $(\exists x)(\forall y)(y \in x)$ *does* have a normal proof!

$$\cfrac{\cfrac{\cfrac{x = x}{x \in \{y : y = y\}}}{(\forall x)(x \in \{y : y = y\})}}{(\exists y)(\forall x)(x \in y)}$$

Nfinite sets are closed under unions of **disjoint** singletons:

$$\text{Nfin} \;=\; \bigcap\{Y : \emptyset \in Y \wedge (\forall xy)(y \notin x \in Y \to x \cup \{y\} \in Y)\}$$

There is an alternative definition of Nfinite that has some motivation and history to it. In the classical setting there is a definition of *natural number* due to Quine (tho' it may go back earlier) as something x s.t. every set of cardinals that contains x and is closed under subtraction-of-1 contains 0. The corresponding definition of finite *set* says that a set is finite (well, *Nfinite*) iff every set that contains it and is closed under **subcision** contains the empty set. Subcision (i learnt this word from Allen Hazen) is the operation $x, y \mapsto x \setminus \{y\}$. Let us forstall all possibility of unclarity by saying that if X is closed under subcision we mean that $x \setminus \{y\} \in X$ for all $x \in X$, never mind where y is.

It might be a useful exercise to show that these two definitions of Nfinite are constructively equivalent. Let's do it.

Remark 2 *The two definitions of Nfinite are equivalent.*

Proof:

Well, the empty set is Nfinite-in-the-new-sense. Suppose X is Nfinite-in-the-new-sense and $x \notin X$. Then every set that contains $X \cup \{x\}$ and is closed under subcision contains X. This is where we use $x \notin X$, since we need $(X \cup \{x\}) \setminus \{x\}$ to be equal to X and the best way to secure that is to require $x \notin X$. And—by induction hypothesis—every set that contains X and is closed under subcision contains $\emptyset$, and that makes $X \cup \{x\}$ Nfinite-in-the-new-sense. So, by induction, every Nfinite set is Nfinite-in-the-new-sense.

For the other direction suppose X is Nfinite-in-the-new-sense. Consider the closure-under-subcision of $\{X\}$. This is an inductively defined set, and we prove using its domestic induction principle that every member of it is $X \setminus Y$ for some Nfinite Y. So $\emptyset$ is of the form $X \setminus Y$ where Y is Nfinite. But then $X = (X \setminus Y) \cup Y$ is a union of two Nfinite sets and is Nfinite.

■

It may be that this second definition gives a more expeditious proof that every Kfinite set is notnot Nfinite. But then again it might not.

A set is **subfinite** if it has a Kfinite superset.

2.1 Some Logical Banalities that may be useful

Readers who are not familiar with constructive logic may not know that constructively

> $\neg\exists$ implies $\forall\neg$ but not conversely
>
> Notice that you have have as many $\exists$ as you like:
>
> $\neg\exists \cdots \exists$ implies $\forall \cdots \forall\neg$
>
> $\exists$ implies $\neg\forall\neg$ but not *vice versa*;

∀ implies ¬∃¬ but not *vice versa*;

¬¬∀ implies ∀¬¬ but not *vice versa*;

∃¬¬ implies ¬¬∃ but not *vice versa*;

¬∀ does not imply anything but ¬∀¬¬ implies ¬¬∃¬ …the point being that ¬¬∀ is stronger than ∀¬¬ so denying the second is strong enough to imply something useful.

Worth checking whether or not ¬¬∨ is associative. I think not

Lemma 3 *Johnstone's weak de Morgan principle [26] is equivalent to the principle that ¬¬ distributes over ∨*

Proof:

Assume $\neg\neg(A \vee B)$. We will deduce $\neg\neg A \vee \neg\neg B$ by using the two following cases of weak de Morgan: $\neg\neg A \vee \neg A$ and $\neg\neg B \vee \neg B$. These two cases give us four possibilities. Three of those four possibilities have either $\neg\neg A$ or $\neg\neg B$, and clearly those three cases will give $\neg\neg A \vee \neg\neg B$. The one remaining case leaves us the chore of deducing $\neg\neg A \vee \neg\neg B$ from $\neg\neg(A \vee B)$, $\neg A$ and $\neg B$. Clearly we deduce $\bot$ and then use *ex falso*.

$$\dfrac{\dfrac{\dfrac{\dfrac{[A]^1 \qquad \neg A}{\bot}\ \to\text{-elim} \qquad \dfrac{[B]^1 \qquad \neg B}{\bot}\ \to\text{-elim} \qquad [A \vee B]^2}{\bot}\ \vee\text{-elim } (1)}{\neg(A \vee B)}\ \to\text{-int } (2) \qquad \neg\neg(A \vee B)}{\bot}\ \to\text{-elim} \tag{1}$$

For the other direction reflect that $\neg\neg(p \vee \neg p)$ is a constructive thesis. If we distribute ¬¬ we obtain PTJ's weak de Morgan.

■

Horn formulæ behave quite well in a constructive setting:

Lemma 4 *Let ϕ be a Horn property and R a relation-in-extension.*
Then $\neg\neg\phi(R) \to \phi(\sim\sim R)$.

Proof:

Let ϕ be the Horn formula

$$(\bigwedge_{i \in I} p_i) \to q$$

where the p_i are atomics of the form '$\langle x, y \rangle \in R$' and q is another such formula or is possibly $\bot$[3]. Now the assertion that the relation R has the property ϕ is

[3] q can't be an equation; antisymmetry is Horn but this lemma doesn't hold for it. We need to put in some more work.

$$(\forall \vec{x})(\bigwedge_{i \in I} p_i \to q)$$

where the quantifier binds all the variables appearing in the p_i and q. So consider the assertion that $\neg\neg(\phi(R))$. This is

$$\neg\neg(\forall \vec{x})(\bigwedge_{i \in I} p_i \to q)$$

We can import the '$\neg\neg$' to infer

$$(\forall \vec{x})\neg\neg((\bigwedge_{i \in I} p_i) \to q)$$

and again (beco's $\neg\neg(A \to B)$ implies $\neg\neg A \to \neg\neg B$))

$$(\forall \vec{x})(\neg\neg(\bigwedge_{i \in I} p_i) \to \neg\neg q)$$

Finally $\neg\neg$ distributes over $\wedge$ to give

$$(\forall \vec{x})((\bigwedge_{i \in I} \neg\neg p_i) \to \neg\neg q)$$

...but this is simply to say that $\sim\sim R$ has ϕ.

■

It may not work for antisymmetry, but on the up-side it will work for things other than relations. The property of being a filter in $\mathcal{P}(V)$ is Horn, so presumably the double complement of a filter is a filter:
Suppose $\neg\neg(A \in F)$ and $A \subseteq B$, well, clearly $\neg\neg(B \in F)$
Suppose $\neg\neg((A \in F) \wedge (B \in F))$. This gives $\neg\neg(A \cap B \in F)$

It seems a suspiciously good fit... $\neg\neg$ distributes over precisely the things it needs to distribute over (namely $\to$ and $\wedge$) to get Horn formulæ to behave well in this context ...but does not distribute over $\vee$. On the other hand I wonder if there is something quite general going on here ...nothing specifically to do with constructive logic ...

Infinite distributivity fails. $A \vee (\forall x)B$ implies $(\forall x)(A \vee B)$ but not conversely. This prevents us from defining a complement x^* of x as $\bigcap\{y : x \cup y = V\}$. There is no reason to expect that $x \cup \bigcap\{y : x \cup y = V\} = V$.

2.1.1 AC

We start with the following observation, which is standard in the literature, but may not be familiar to NFistes.

Remark 5 *(Diaconescu [15]) AC $\to$Law of excluded middle.*

Proof:

Take AC in the form that every inhabited set of inhabited sets has a choice function.

Let $a = \{x : x = \bot \vee (p \ \wedge \ x = \top)\}$ and
$b = \{x : x = \top \vee (p \ \wedge \ x = \bot)\}$ and
$X = \{a, b\}$.

Then X is an inhabited set of inhabited sets and must have a selection function f. Therefore $f(a) \in a$ and $f(b) \in b$. Further we know $(f(a) = \bot) \vee (f(a) = \top)$ and $f(b)$ similarly. Thus there are four possibilities, so we can use proof by cases. If $f(a) = \top$ then p; if $f(b) = \bot$ then p. If neither of these happens—so $f(a) = \bot$ and $f(b) = \top$—then at least $f(a) \neq f(b)$ so $a \neq b$. But $p \to (a = b)$, so we infer $\neg p$. Thus proof by cases gives us excluded middle.

■

Of course in the NF context this proof establishes only that that form of AC implies excluded middle for *weakly stratified* formulæ.

3 Why *i*NF should be weak

There is a standard trick in classical propositional logic for taking a theory T and a theorem A of T and axiomatising T in such a way that A is an independent axiom of the new axiomatisation. The idea is that one can then consistently replace A by $\neg A$ and maybe something interesting will happen.

I am going to need a constructive version of this result, so i'll spell it out.

Let T and A be as above. We also assume (for the avoidance of triviality) that $\neg\neg A$ is not a thesis of constructive first-order logic.

We axiomatise T with the axiom A and the scheme which we write '$\Sigma_B A \to B$', of conditionals $A \to B$, where the Bs are the theorems of T. Suppose A is not independent, and that it follows from the scheme $\Sigma_B A \to B$. Then, for some finite set of Bs, we have

$$\bigwedge_B (A \to B) \vdash A$$

where the turnstile means provability in constructive first-order logic. But $\bigwedge_B (A \to B)$ is equivalent to $A \to \bigwedge_B B$ whence

$$(A \to \bigwedge_B B) \vdash A.$$

and

$$\vdash (A \to \bigwedge_B B) \to A.$$

Now

$$\vdash ((A \to \bigwedge_B B) \to A) \to \neg\neg A$$

so so (by *modus ponens*)

$$\vdash \neg\neg A,$$

contradicting assumption. So A is an independent axiom.

■

The idea is to do this in a constructive setting where we replace A by $\neg A$ as allowed, but retain as an axiom something *weaker* than A. To this end we say that a formula in $\mathcal{L}(T)$ that is classically equivalent to A but does not imply $\neg\neg A$ constructively is a *weakening* of A, and we will denote such a formula by '$w(A)$'. Observe that Glivenko's thereom that in propositional logic the double negation of a classical tautology is intuitionistically correct means that in propositional logic any $w(A)$ that is classically equivalent to A must constructively imply $\neg\neg A$, so such weakeners are not to be had in propositional logic; their construction must involve manipulation of the quantifiers. We are obliged to work in first-order logic

Now let us say that a closed formula A is *sufficiently complex* if it admits such a weakening.

In particular if A is sufficiently complex it is stronger than its double negation. Notice that anything of the form $\neg A$ stands a good chance of being sufficiently complex, since it is equivalent to its double negation. Negations of ambiguity axioms are almost guaranteed to be sufficiently complex. But I'm getting ahead of myself.

Let T be a theory, A a sufficiently complex theorem of it. As before we assume (for the avoidance of triviality) that $\neg\neg A$ is not a thesis of constructive first-order logic.

We axiomatise T by A and the scheme $\Sigma_B A \to B$, where the Bs are the theorems of T, as above. So we can replace A in this axiomatisation by $\neg A$. That much is standard. The novel point here is that one can do that while retaining $w(A)$ as an axiom! The proof parallels the foregoing.

Suppose *per contra* that

$$w(A), \Sigma_B A \to B \vdash \neg\neg A.$$

By compactness we need only finitely many of these conditionals on the left, and—as above—we recall that, for any fixed A, any finite conjunction of things of the form $A \to B$ is another thing of that form. Thus we get

$$w(A), A \to \bigwedge B \vdash \neg\neg A$$

for some finite set of Bs whence

$$w(A) \vdash (A \to \bigwedge B) \to \neg\neg A.$$

But now we recall (as above) that, for any X,

$$\vdash ((A \to X) \to \neg\neg A) \to \neg\neg A$$

whence

$$w(A) \vdash \neg\neg A$$

But this contradicts the assumption that A was sufficiently complex.

$\blacksquare$

This gives us an operation on theories. Input a consistent theory T, and A a sufficiently complex theorem of T. Output a (constructive) theory T^w which has axioms: $w(A)$, $A \to B$ for all axioms B of T ... **and** $\neg A$. **And** T^w is consistent!

OK, i admit it; this is a trivial factoid. So far it only says that you can "turn" (Sorry, i have been watching too much Le Carré of late) a single theorem. Naturally one would like to be able to turn infinitely many expressions simultaneously. I am pretty sure that that is possible but I haven't yet worked out the best way to do it.

I don't think i am letting any secrets out of the bag by saying that the medium term project is to tackle a *strongly typed* theory (as it might be T$\mathbb{Z}$T). Suppose T$\mathbb{Z}$T proves some negation of an ambiguity axiom. (I don't suppose for a moment that it does—i have far more faith in Randall than that—but in principle it might). Then you 'turn' all the negations of instances (or disjunctions of such negations) of the ambiguity scheme that T$\mathbb{Z}$T proves. (All such formulæ are sufficiently complex). Then you have a theory that is classically equivalent to T$\mathbb{Z}$T and consistent with the ambiguity scheme, so you add the scheme and then *you can drop the type indices*. This doesn't neccessarily give you *i*NF but it does give you a constructive theory in the language of set theory whose axioms are classically equivalent to it. The interest in the foregoing (if any) lies in the fact that it parallels the *aperçu* that realizability gives us a reason to believe in the consistency of *i*NF that is not at the same time a reason to believe in the consistency of NF. It's a straw in the wind.

This has been at the back of my mind for some time, and is the cause of my firmly-held suspicion that *i*NF is weak.

Two things to think about:

(i) How do we 'turn' infinitely many formulæ simultaneously?

(ii) Even if we can, the version of sl iNF that we get might be too weak.

(iii) Does this invite any observations about realizability.

Let's have a go at (i). Let T be a classical theory, and $\mathcal{A}$ a set of theorems of T which we wish to turn. We need $\overline{\mathcal{A}} = \{\neg A : A \in \mathcal{A}\}$ to be a consistent theory (o/w we aren't going to be able to turn $\mathcal{A}$!) What do we have to add to $\mathcal{A}$ to axiomatise T? The things we add have to be theorems of T and they have to be consistent with $\overline{\mathcal{A}}$ and the aggregate of them has to axiomatise T. Call this set we are trying to build '$\mathcal{B}$'. Why not just let B be the set of theorems of T that are consistent with $\overline{\mathcal{A}}$. The problem is that the set of such theorems is not closed under conjunction.

We want X a $\subseteq$-maximal subset of T with the property that $X \cup \overline{\mathcal{A}}$ is consistent. The set of such sets is a chain-complete poset so it will have a maximal element. But will it have a maximal element that axiomatises T?

No reason why it should. But we can at least do the following. Suppose there are finitely many things that we wish to turn. Then consider the *disjunction* of them. Do the above trick, then you get a model that satisfies the weakening of the disjunction, plus all the negations. Can't we fiddle things so it satisfies the weakening of the *con*junction.

4 The Constant Domain principle

$$(\forall x)(C \vee A(x)) \to .C \vee (\forall x)A(x)$$

This is something we need to look at very closely. Does it imply NF?

There is a result of Grischa's that says that interpolation fails.

5 Fishy Sets

We say of two variables 'x' and 'y' that 'x' is connected to 'y' if there is an atomic formula containing 'x' in which 'y' or some variable connected to it occurs. A formula is **Crabbé-elementary** if for every quantifier, the only variables occurring in its scope are variables connected to the variable bound by that quantifier.

(Classically every formula is equivalent to a Crabbé-elementary formula. Intuitionistically it is not the case that every formula is equivalent to one that is Crabbé-elementary, and this makes the intuitionistic case much more complex.

[*HOLE At this point one should insert some brief clarification of distributive laws and an itemisation of those laws whose constructive failure is implicated in the existence of fishy sets... or (perhaps!) an explanation of why it's nothing to do with distributive laws which—now i start to think of it—it appears not to be.*]

The *fishy sets* involved in the deduction of excluded middle from constructively questionable principles all make essential use of formulæ that are not Crabbé-elementary.) This leads us to a definition.

Definition 6 *For all a and b, and for all $p \in \Omega$, the set*

$$\{x : ((x \in a) \wedge p) \vee ((x \in b) \wedge \neg p)\}$$

is a **fishy combination**[4] *of a and b.*

[4] I learnt the word from Douglas Bridges (tho' he says he in turn learnt it from Ian Stewart).

Classically a fishy set is a definable set that is identical to *one* of two (or perhaps more) things (but you don't know which) and it's distinct from the other. Constructively the set with the same definition is not actually distinct from both of them.

Let's spell this out.

REMARK 7 *Suppose $a \neq b$.*
Then the fishy combination:

$$\texttt{fishy} = \{x : ((x \in a) \wedge p) \vee ((x \in b) \wedge \neg p)\}$$

is not distinct from both a and b.

Proof:
Clearly $p \to \texttt{fishy} = a$, whence $\texttt{fishy} \neq a \to \neg p$, and analogously $\neg p \to \texttt{fishy} = b$, whence $\texttt{fishy} \neq b \to \neg\neg p$. So if $\texttt{fishy}$ is distinct from both a and b we infer $\neg p \wedge \neg\neg p$, which is impossible.

■

The challenge is to get a good notion of fishy combination of more than two things. The fishy combination of two things relied on two two values the conjunction of whose negations was $\perp$. Clearly we need three truth-values the conjunction of whose negations is $\perp$.

Deep breath.

Consider a discrete set A, with $t \in {\sim\sim}A$. (Any two things in A are equal or unequal). Consider the fishy combination

$$\texttt{fishy} =: \{x : (\exists a \in A)(x \in a \wedge a = t)\}$$

Let a be an arbitrary member of A. If $a = t$ then $\texttt{fishy} = a$. So if $\texttt{fishy} \neq a$ we must have $a \neq t$. So if $\texttt{fishy}$ is not equal to any member of A we must have t not equal to any member of a. So $t \notin A$. But $t \in {\sim\sim}A$. So $\texttt{fishy}$ cannot be distinct from all members of A. But then it is in ${\sim\sim}A$.

But of course what we really want is that $\texttt{fishy} = t$.

REMARK 8 *For any two sets $a \neq b$ there are sets a' and b' such that*

$$a' \neq b',\ \neg(a' \neq a \wedge a' \neq b) \text{ and } \neg(b' \neq a \wedge b' \neq b).$$

[we need to restate this carefully, since we could have $a = a'$ and $b = b'$!]

Proof:
Fix p for the moment (tho' we can of course vary it). Given a and b form $a' = \{x : ((x \in a) \wedge p) \vee ((x \in b) \wedge \neg p)\}$ and $b' = \{x : ((x \in b)) \vee ((x \in a) \wedge \neg p)\}$.

Let's check that neither of a' or b' can be distinct from both a and b.

If p then $a' = a$ and $b' = b$;

If $\neg p$ then $a' = b$ and $b' = a$.

Since we have $\neg\neg(p \vee \neg p)$ we infer

$$\neg\neg((a' = a \wedge b' = b) \vee (a' = b \wedge b' = a))$$

which implies

$$\neg[\neg(a' = a \wedge b' = b) \wedge \neg(a' = b \wedge b' = a)].$$

Now suppose *per impossibile* that both $a' \neq a$ and $a' \neq b$. Then both $\neg(a' = a \wedge b' = b)$ and $\neg(a' = b \wedge b' = a)$ can be simplified to $\top$! So we infer the false. So a' cannot be distinct from both a and b. *Mutatis mutandis* neither can b' be distinct from both a and b.

$\blacksquare$

Theorem 9

For any two sets a and b the function defined on Ω by

$$p \in \Omega \;\mapsto\; \{x : ((x \in a) \wedge p) \vee ((x \in b) \wedge \neg p)\}$$

is injective.

Proof:

This function is evidently total; it remains to be shown that it is injective. Suppose

$$\{x : ((x \in a) \wedge p) \vee ((x \in b) \wedge \neg p)\} = \{x : ((x \in a) \wedge q) \vee ((x \in b) \wedge \neg q)\}.$$

This is the same as

$$(\forall x)((((x \in a) \wedge p) \vee ((x \in b) \wedge \neg p)) \longleftrightarrow (((x \in a) \wedge q) \vee ((x \in b) \wedge \neg q))).$$

Fix x. Then

$$(((x \in a) \wedge p) \vee ((x \in b) \wedge \neg p)) \longleftrightarrow (((x \in a) \wedge q) \vee ((x \in b) \wedge \neg q)).$$

Assume p and the LHS. Then $x \in a$ so, by using the L $\to$ R implication, we infer $((x \in a) \wedge q \;\vee\; (x \in b) \wedge \neg q)$. We can't have the second disjunct co's that would imply $x \in b \wedge \neg p$, contradicting assumption. So we must have the first disjunct, giving q. Thus $p \to q$. The other direction is analogous. $\blacksquare$

I have found thinking about fishy sets to be quite helpful. The double complement operator $\sim\sim$is inflationary, order-preserving and idempotent, but it isn't very continuous. It preserves meets ($\cap$)—indeed kfinite meets—but not joins, and not arbitrary intersections. $\sim\sim\{a\}$ is precisely the set of things notnotequal to a; but $\sim\sim\{a, b\}$ contains not just things either notnotequal to a or notnotequal to b; it also contains all fishy combinations of a and b. Is the converse inclusion correct?

Watch this space.

Suppose $x \in \sim\sim\{a, b\}$. If $x \neq a$ and $x \neq b$ then $\neg(x = a \vee x = b)$, whence $x \notin \{a, b\}$. So if $x \in \sim\sim\{a, b\}$ it cannot be distinct from both a and b. We seek a proposition p such that x is a fishy combination of a and b using p. p could be something like $a = x$.

$$x = \{u : (u \in a \wedge (x = a)) \vee (u \in b \wedge (x \neq a))\}$$

We can define a'' as $\{x : x \in a' \wedge p. \vee .x \in b' \wedge \neg p\}$.

But $x \in a' \wedge p$ simplifies to $x \in a \wedge p$, and $x \in b' \wedge \neg p$ simplifies to $x \in a \wedge \neg p$. So a'' turns out to be $\{x : x \in a \wedge (p \vee \neg p)\}$.

This doesn't actually show that every set is fishy, but the warnings are clear enough for all to see. If $a = \sim\sim a$ then $\sim\sim a'' = a$, so every stable set is the double complement of a fishy set. That's enough to put us on notice that fishy sets may be everywhere, poisoning wells, molesting our daughters.... We need to be on our guard.

Annoying they may be, but fishy sets are quite useful, as we are about to see.

Lemma 10 $\sim\sim A$ *contains all fishy combinations of members of* A*.*

Proof:

Suppose a and b are both in A. Let c be a fishy combination of a and b. If $c \notin A$ then clearly $c \neq a$ and $c \neq b$. But c is a fishy combination of a and b so we have $\neg(c \neq a \wedge c \neq b)$. So $\neg(c \notin A)$ which is to say $c \in \sim\sim A$. ■

Is there anything like a converse to lemma **??**? Is everything in $\sim\sim A$ a fishy combination of things in A?

Lemma 11 *Suppose* τ *is a permutation that moves a set* a*. Then* τ *moves every fishy combination of* a *and* $\tau(a)$*.*

Proof: Suppose $\tau(a) = b$, and c is a fishy combination of a and b such that $\neg(a \neq c \wedge b \neq c)$. If τ is a permutation that fixes c then $c \neq a$ (beco's a is moved and c is fixed) and $c \neq b$ similarly. But we deny this conjunction, whence c is moved. ■

The transposition (a, b) swaps a and b and fixes everything else. For this transposition to exist every x must be equal to one of a and b or distinct from both. But this can never happen, beco's a and b have fishy combinations which (a, b) doesn't know what to do with.

So any permutation that moves anything moves quite a lot of things. Can we be more specific? Can we tie down '*quite a lot*'?

Theorem 12 *Suppose* τ *is a permutation and, for some* a*,* $\tau(a) \neq a$*.*
Then $\{x : \tau(x) \neq x\}$ *is not Kfinite (unless the logic is classical).*

Proof: The idea is that we map $\{x : \tau(x) \neq x\}$ onto Ω, the truth-value algebra, and then appeal to the two facts (both proved elsewhere in these notes) that (i) Ω is not kfinite unless the logic is classical (corollary 47), and (ii) a surjective image of a Kfinite set is Kfinite (lemma 50).

So: let us map $\{x : \tau(x) \neq x\}$ onto Ω.

Fix some a such that τ moves a. We will define a surjection $f : \{x : x \neq \tau(x)\} \twoheadrightarrow \Omega$. Declare that f sends c to $\{x : x = \emptyset \wedge c = a\}$. Now let p be an arbitrary member of Ω. We seek c s.t. $f(c) = p$. The obvious candidate is $c = \{x : x \in a \wedge (\{\emptyset\} = p))\}$. If c is fixed then $c \neq a$ so $p = \bot$. c will be sent to

$$\{x : x = \emptyset \wedge (\{x : x \in a \wedge (\{\emptyset\} = p))\} = a)\}.$$

which simplifies to p as follows.

The displayed set abstract clearly points to a subset of $\{\emptyset\}$. This subset will be p as long as the boolean inside the parenthesis, namely $(\{x : x \in a \wedge (\{\emptyset\} = p))\} = a)$ simplifies to (the proposition) p, which is of course $p = \{\emptyset\}$. This boolean is an equation, and the LHS is clearly a subset of the RHS, with equality as long as the boolean after the '$\wedge$' is true. So the boolean inside the parenthesis is just p.

■

This is as much as to say that $\sim\sim X$ contains all fishy combinations of members of $\sim\sim X$.

[What about a fishy combination of a with a? It's notnotequal to a! Is everything notnotequal to a a fishy combination of something with itself?]

Fishy sets help us prove the following

Remark 13 *No stable set with two distinct members can be kfinite unless the logic is classical.*

Proof: Let X be stable, with distinct members a and b. We map X onto Ω by sending each $x \in X$ to $[[x = a]]$. Let p be any truth-value. Consider the fishy combination $\{y : (y \in a \wedge p) \vee (y \in b \wedge \neg p)\}$ (or something like that!) This is equal to a with truth-value p, and p was an arbitrary member of Ω so our map is onto Ω.

5.1 Fishy Sets show there are no Isolated Sets

Let us—for the moment (and it will only be for the moment since i propose to show that there aren't any)—say that a set a is *isolated* if $(\forall x)(x = a \vee x \neq a)$.

Theorem 14 *If there are any isolated sets then the logic is classical.*

Proof: Suppose a is isolated, and $b \neq a$. Let p be an arbitrary proposition and consider the fishy set $c = \{x : (x \in a \wedge p) \vee (x \in b \wedge \neg p)\}$. Since a is isolated we have $c = a$ or $c \neq a$. If $c = a$ then p follows. If $c \neq a$, then—by fishiness—c cannot be distinct from b. This gives $\neg\neg\neg p$ which is of course $\neg p$. ■

Corollary 15 *If there are any simple transpositions then the logic is classical.*

Proof:

Suppose there is a permutation τ that swaps a and b and fixes everything else. That is to say $a \neq b$ and $(\forall x)((x = a \wedge \tau(x) = b) \vee (x = b \wedge \tau(x) = a) \vee \tau(x) = x)$.

Let x be arbitrary. Then either $x = a$, or $x = b$ (in which case $x \neq a$), or $\tau(x) = x$, in which case—again—$x \neq a$. So a is isolated, contradicting theorem 14. ■

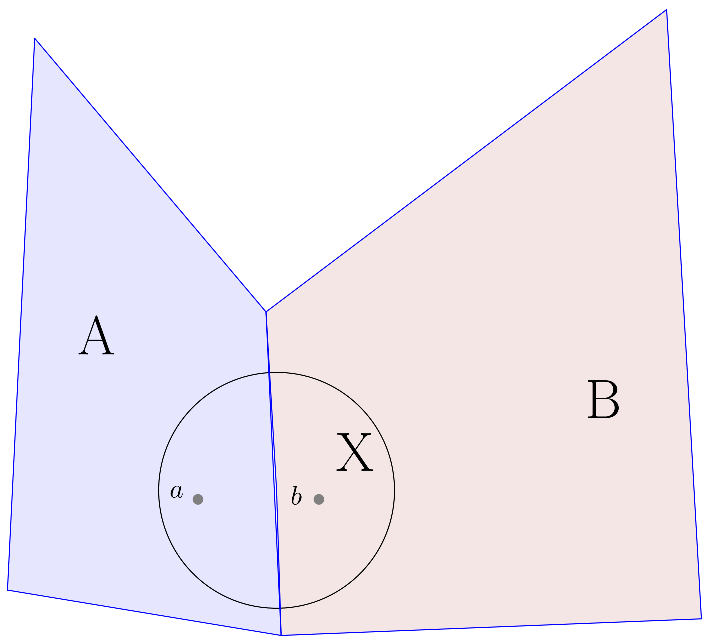
5.2 Fishy Sets Constrain the Nature of Partitions of Stable Sets

Recall that a set X is stable if $(\forall y)(\neg\neg y \in X \to y \in X)$.

Lemma 16 *Let X be a stable set. Then there is no pair of sets A, B—both of which meet X—such that*

$$(\forall x \in X)((x \in A \vee x \in B) \wedge (x \notin A \vee x \notin B))$$

...unless the logic is nearly classical, of course.



Where is c??

Proof: Well, suppose there is such a pair A, B. Take $a \in A \cap X$ and $b \in B \cap X$ and consider a fishy combination $c = \{x : ((x \in a) \wedge p) \vee ((x \in b) \wedge \neg p)\}$.

We have $\neg\neg(c \in X)$ by lemma 10, but X is stable, whence $c \in X$.

Now ask: "To which of A and B does c belong?"

If $c \in B$ then it cannot be equal to a. But it cannot be distinct from *both* a and b so we have $\neg\neg(c = b)$. But $c = b$ implies $\neg p$ whence $\neg\neg(c = b)$ implies $\neg\neg\neg p$ which implies $\neg p$. Analogously if $c \in A$ we infer $\neg\neg p$. So we infer $\neg\neg p \vee \neg p$, for arbitrary p. ■

Notice that we do really need the set we are putatively partitioning into precisely two pieces to be stable.

But i think we can do even better. I claim that

Theorem 17
Let X be stable. Then X has no partition into Nfinitely many pieces unless the logic is (nearly) classical—$\neg\neg p \vee \neg p$.

Proof:
Fix X; we prove by induction on Nfinite sets that for no Nfinite set A is there $y \notin A$ such that $A \cup \{y\}$ is a partition of X.

By lemma 16 this works for singletons. Suppose true for A; we want it to be true for $A \cup \{x\}$, where $x \notin A$. Suppose there is $y \notin (A \cup \{x\})$ such that $A \cup \{x\} \cup \{y\}$ is a partition of X. But then $A \cup \{x \cup y\}$ is a partition of X, contradicting induction hypothesis on A.

Ah! But we need $\{x \cup y\} \notin A$ ■

Have to be very careful how we state this ... On hold for the moment

You *want* to say: lemma 16 says that (unless the logic is nearly classical) you can split a stable set into precicely two pieces. So (you want to say) if you can't split it into n (n is Nfinite) pieces then you can't split it into $n+1$ pieces, co's if you could you could merge two of those pieces and get a partition into n pieces. But you have to say it without using numbers. Should be do-able tho' ...

Show by induction that the sumset of a kfinite partition is never stable. But that would need the family of kfinite partitions to be defined by breaking pieces into two

Corollary 18 *The quotient $V/(\neg\neg =)$ is not kfinite.*

Proof: This is beco's $V/(\neg\neg =)$ maps onto every partition Π_X, and any such partition in turn maps onto Ω.

[Should write out a very clear proof of this.]

`\begin{insert}`

Given a truth-value $a \subseteq \{\emptyset\}$, seek x such that $[[x = \emptyset]] = a$.

Now $[[x = \emptyset]] = \{y : y = \emptyset \wedge x = \emptyset\}$, so we seek x s.t.

$\{y : y = \emptyset \wedge x = \emptyset\} = x$.

The obvious thing to try [this was what you were thinking, isn't it] is $x = \bigcup a$. So we want

$a = \{y : y = \emptyset \wedge \bigcup a = \emptyset\}$

Isn't this just the same as $a = \{\bigcup a\}$? And isn't that true? Perhaps not... $\bigcup a = \emptyset$

`\end{insert}`

This has the makings of another proof. Sse $V/(\neg\neg =)$ is kfinite. Then it has a transversal, T. Then $(\forall x)(\exists! y \in T)(\neg\neg(x = y))$. Let a, b be two distinct members of T, and p any proposition. Consider the fishy combination $f_p = \{z : z \in a \wedge p. \vee z \in b \wedge \neg p\}$. There is $x \in T$ such that $\neg\neg(x = f_p)$. If $x \in T$ is notnotequal to f_p it cannot be distinct from both a and b so we get $\neg\neg(\neg\neg(f_p = a) \vee \neg\neg(f_p = b))$ ...

Eurgh

Remark 19 *The quotient $V/(\neg\neg =)$ can be partitioned into pairs.*

Proof:
Observe that x and $V \setminus x$ are always distinct. Observe, too, that if $\neg\neg(x = y)$ then $\neg\neg((V \setminus x) = (V \setminus y))$—and in fact $(V \setminus x) = (V \setminus y)$. Suppose $z \notin x$, and $\neg\neg(x = y)$; then $z \notin y$ so by extensionality .... The partition of $V/(\neg\neg =)$ into pairs is $\{\{[x]_{\neg\neg=}, [V \setminus x]_{\neg\neg=}\} : x \in V\}$ ■

It would be nice to be able to show that if we remove one element from it it can still be partitioned into pairs.

We knew that there was a set—V—that was not kfinite; however V is not discrete. Here we have an example of a discrete set that is not kfinite, and that is a stronger result. Unfortunately this doesn't seem to be *quite* strong enough to prove that *i*NF interprets Heyting Arithmetic.

Of course any set that maps onto a stable set (or set lacking a small partition) also has no small partition.

One obvious natural partition of V that is of some interest is the collection of orbits of Symm(V). We have the following immediate corollary of theorem 17:

Corollary 20 *The collection of orbits of Symm(V) is either a singleton or is not kfinite.*

It would be nice to know which of these possibilities is the case....

Actually we need to be careful—the last three claimed results in this section are not secure.

Still thinking about the possibility of a Nfinite set x s.t $\sim\sim x = V$. Life is complicated by the fact that $\neg\neg(y \in x)$ does not imply that $(\exists y' \in x)(\neg\neg(y = y'))$. But what if x is Nfinite?

Suppose $\neg\neg(x \in A \cup \{a\})$ but $x \notin A$. Then it's easy to see that $\neg\neg(x = a)$. Can this power an induction? Clearly not, but might it be the case that everything in $\sim\sim (A \cup \{a\})$ is a fishy combination of a with something in (notnotin?) A?

It would be nice if we could prove that everything notnotin $A \cup \{a\}$ is a fishy combination of a and something in (notnotin?) A. Should start with everything notnotin $\{a, b\}$.

5.3 Fishy sets and PTJ Weak de Morgan Principle

The weak de Morgan principle is $\neg p \vee \neg\neg p$.

Suppose a and b are two distinct sets, and let c be a fishy combination of a and b. Assume weak de Morgan.

Then we have

$$(\neg\neg(c = a) \vee (c \neq a)) \wedge (\neg\neg(c = b) \vee (c \neq b))$$

Distributing we get four disjuncts, one of which is $c \neq b \wedge c \neq a$ which we can discard. We can also discard $\neg\neg(c = a) \wedge \neg\neg(c = b)$ since that implies $\neg\neg(a = b)$. There remain two:

$$\neg\neg(c = a) \wedge c \neq b \text{ and } \neg\neg(c = b) \wedge c \neq a$$

which gives

$$\neg\neg(c = a) \vee \neg\neg(c = b)$$

Humph.

Well, at least it shows that weak de Morgan for atomic formulæ implies weak de Morgan for stratified formulæ

The obstacle to interpreting HA in iNF has always seemed (to me) to be the possibility of an Nfinite set that is *dense*: its double complement is V. I have never found a way of refuting this possibility. I now think i have found such a way. Bear with me ...

Suppose $\mathcal{V}$ is a dense Nfinite set, with $a \neq b$ both $\in \mathcal{V}$. Let p be a truth value and consider $c_p = \{x : (x \in a \wedge p) \vee (x \in b \wedge \neg p)\}$. (The propositional variable 'p' is of course syntactic sugar for a set variable ranging over inhabitants of Ω, the power set of $\{\emptyset\}$, so that 'p' (as in the definition of c_p above) is short for '$p = \{\emptyset\}$'). If $c_p \in \mathcal{V}$ we have $c_p = a \vee c_p \neq a$ and ${}_pc = b \vee c_p \neq b$. Distributing we get $c_p = a \vee c_p = b$ (since $a \neq c_p \neq b$ is not possible) which implies $p \vee \neg p$. So we have

$$(\forall p)(c_p \in \mathcal{V} \rightarrow (p \vee \neg p))$$

So sse $(\forall p)(c_p \in \mathcal{V})$. Then classical logic holds for weakly stratified expressions. So, if it doesn't, we infer $\neg(\forall p)(c_p \in \mathcal{V})$. But that isn't enough to give a contradiction.

But suppose it had worked. Suppose there is no maximal Nfinite set. That is, for X Nfinite, we have $\neg(\forall y)\neg\neg(y \in X)$, and this implies that, for all Nfinite X, $\neg\neg(\exists y)(y \notin X)$. That is to say: we have $(\forall n \in \mathbb{N})\neg\neg(\exists m)(m = n + 1)$.

Suppose $\mathbb{N}$ is finite. Then we can use Linton-Johnstone to infer $\neg\neg(\forall n \in \mathbb{N})(\exists m)(m = n + 1)$.

Now $(\forall n \in \mathbb{N})(\exists m)(m = n + 1)$ implies that $\mathbb{N}$ is not finite. So its double negation has the same (negative) consequence. So $\mathbb{N}$ is not finite.

Is that enough? I think so. It gives us an interpretation of HA on the nose. The naturals of the interpretation of HA are just the cardinals of the Nfinite sets.

Lemma 21 *Let $\mathcal{V}$ be a dense set. Then*

$$(\exists x)\phi \rightarrow \neg\neg(\exists x \in \mathcal{V})\phi$$

Proof:

Assume $\neg(\exists x \in \mathcal{V})\phi$, and that x satisfies ϕ. Then x cannot be in $\mathcal{V}$, beco's by assumption $\mathcal{V}$ does not contain anything which is ϕ. But $\neg\neg(x \in \mathcal{V})$ So $\neg\neg(\exists x \in \mathcal{V})\phi$ after all. ■

Corollary 22 *Let $\mathcal{V}$ be a dense set. Then*

$$\neg\neg(\exists x)\phi \longleftrightarrow \neg\neg(\exists x \in \mathcal{V})\phi$$

Proof:
Easy-peasy
We can do something similar for $\forall$

Remark 23 *Let $\mathcal{V}$ be dense. Then*
$\forall x \neg\neg\phi \longleftrightarrow \forall x \in \mathcal{V} \neg\neg\phi$

Michael has an argument that shows that if x and y are both Nfinite and are not-not-equal then they have the same cardinal. (13/vii/20, californian time):

> "If x is a finite set then $Nc(x)$ is a finite cardinal. If x and y are two finite sets then $Nc(x) = Nc(y)$ if and only if $x \sim y$. The set F of finite cardinals has decidable equality, so we can double-negate that equivalence and drop the double negation on the left. Thus $Nc(x) = Nc(y)$ if and only if not not $(x \sim y)$. "

The point is that if x and y are two finite sets, we ask about their cardinals, $|x|$ and $|y|$. These two objects are equal or unequal. If they are unequal then $x \neq y$—which will contradict the assumption that $\neg\neg(x = y)$. So we must have the other horn.

This means that any two candidates for $\mathcal{V}$-ness have the same cardinal, since they are notnotequal.

And i think we have a converse: anything that has the same cardinal as $\mathcal{V}$ must also be a $\mathcal{V}$. Suppose $|\mathcal{U}| = |\mathcal{V}|$ but that poor old little x (for excluded) is not a member of $\mathcal{U}$. Then $|\mathcal{U}| + 1$ is defined, and $|\mathcal{V}|$ was not the last cardinal after all. Contradiction. So there is no such x, so $\sim\sim\mathcal{U} = V$.

It would be good to prove that $\mathcal{V}$ is *actually* unique, not just $\neg\neg$ =-unique. The following lemma would be useful:

Suppose x and y are both Nfinite and both the same size as $x \cap y$; then $x = y$.

That would do it. You'd probably need to do a clever induction on $\mathbb{N}$. True for $n = 1$!

Michael says:

> "I guess you mean by "same size" that they are similar, i.e. have the same Nc.
>
> So x and y are finite and both are similar to each other and to $x \cap y$. A member of a finite cardinal is finite, so $x \cap y$ is also finite. If $y \subseteq x$ and x, y finite then $y \setminus x$ is finite. Hence $x \setminus (x \cap y) = x \setminus y$ is finite.
>
> A finite subset of a finite set is a separable subset, so $x = y \cup (x \setminus y)$. Now $x \setminus y$ is empty, since if not, then x is similar to its proper subset $x \cap y$, so x is Dedekind infinite, but a Dedekind infinite set is

not finite, contradicting the assumption that x is finite. Therefore $x = y \cup (x \setminus y) = y \cup \emptyset = y$. Thus $x = y$ as desired. All the lemmas mentioned here are proved in most of my earlier drafts except the lemma that if $y \subseteq x$ and x, y finite then $x \setminus y$ is finite, which is easily proved by induction on finite sets x. "

Observe that $\in$ between members of $\mathcal{V}$ can be taken to be stable. It is certainly true that

$$(\forall xy \in \mathcal{V})\neg\neg(\neg\neg x \in y \to x \in y)$$

so, by Linton-Johnstone, we infer

$$\neg\neg(\forall xy \in \mathcal{V})(\neg\neg x \in y \to x \in y)$$

so we can consistently add

$$(\forall xy \in \mathcal{V})(\neg\neg x \in y \to x \in y).$$

Don't know if this helps.

Actually we can do something more general. After all, we always have $\neg\neg(\neg\neg p \to p$ so—for any stratified ϕ—we get the consistency of $(\forall \vec{x} \in \mathcal{V})(\neg\neg\phi(\vec{x}) \to \phi(\vec{x}))$

We want $\mathcal{V}$ to obey extensionality. So sse a and b satisfy $(\forall xin\mathcal{V})(x \in a \longleftrightarrow x \in b)$. So $(\forall x)(\neg\neg(x \in \mathcal{V}) \to \neg\neg(x \in a \longleftrightarrow x \in b)$.

But everything satisfies the antecedent so we infer
$(\forall x)(\neg\neg(x \in a \longleftrightarrow x \in b)$.

6 Partitions, Permutations and Excluded Middle

Does V have any nontrivial separable subsets at all?

Is it not the case that the *support* of a permutation of a set X is always a separable subset of X?

$(\forall x)\neg\neg(x = f(x) \vee x \neq f(x))$ is true but that not enuff.

Another useful fact about separability. If the domain of f is separable then there is an extension of f to the whole of V.

"Here is the situation. I have made several mistakes in the last month. They are all as it turns out variations on the same mistake. The one we discussed yesterday;

A second example concerns the law that says $2^(x^+) = 2^x + 2^x$, namely, if the right hand side is in F then the left-hand side is too. F is the set of finite cardinals, by definition the least set closed under inhabited successor, i.e. $x \in F$ and x^+ inhabited $\to x^+ \in F$.

Constructively one uses $SSC(x) =$ the set of separable subsets of x, rather than $SC(x)$. One can prove x finite $\to SSC(x)$ finite. Exponentiation is $2^x = Nc(SSC(a))$ where $USC(a) \in x$.

FINITE is the least class containing empty and closed under adding $x \cup \{c\}$ where $c \notin x$.

Addition and order are defined as Rosser and Specker do except using "separable subset" instead of "subset".

Now suppose m is the "maximum integer", i.e. $m \in F$ and $m^+ \notin F$. Then if $U \in m$, U is "unenlargeable", i.e. you can't find $c \notin U$ (else m^+ would exist). Define $\kappa = Nc(SSC(U))$.

Like every other finite cardinal, $\kappa \leq m$. Now the crucial lemma that is really the only thing I haven't proved (everything else has been checked in Lean, but I don't even have a paper proof of this lemma) is

$$\kappa + \kappa \notin F$$

This says that $SSC(U)$ is too big to be copied. $\kappa + kappa \in F$ if and only if there exist two disjoint sets a and b each similar to $SSC(U)$ (then $a \cup b \in \kappa + \kappa$).

Normally if you want to copy $SSC(A)$, you pick some set c that is not in A and your copy is $\{x \cup \{c\} : x \in A\}$. But picking c is just what you can't do in case $A = U$. That of course is not a proof, as perhaps there is some cleverer way to copy $SSC(U)$.

We have $\kappa = 2^{Tm}$ by the definitions of exponentiation and T; check that to see if you understand the definitions.

I have proved "finite DNS": I can move a double-negation out from a "$\forall t \in x$" when $x \in$ FINITE. It follows that every finite set is not-not a subset of U.

I will give it a serious try to prove the lemma stated above; if I can't, and if I still think it should be true, I'll just give what I've done to you and Randall. It has been a fun pandemic project but I'm about out of steam."

7 My reply, work in progress

OK, $\mathfrak{m}$-for-Michael is the largest Nfinite cardinal. Michael wants $2^{T\mathfrak{m}} + 2^{T\mathfrak{m}}$ to be greater than $\mathfrak{m}$. Trouble is, it ain't. We might have to copy everything down through a lens of T so we have room to work, One way of doing that is to repurpose the mathfrak character '$\mathfrak{m}$' to be not Michael's largest Nfinite cardinal number but T of it. Then copy back by T^{-1} at the end.

If Michael's lemma were true, we would have

$$T\mathfrak{m} < 2^{T\mathfrak{m}} < \mathfrak{m} < 2^{T\mathfrak{m}} + 2^{T\mathfrak{m}} = 2^{T\mathfrak{m}+1}$$

in particular

$$2^{T\mathfrak{m}} < \mathfrak{m} < 2^{T\mathfrak{m}+1}$$

The question to ask is: "how many bits are there in the binary representation of $\mathfrak{m}$? Clearly it must be $T\mathfrak{m}$. Now we consider the pressing-down function $f : \mathbb{N} \to \mathbb{N}$ that sends n to the number of bits in the binary representation of n (aka the smallest k s.t $n \leq 2^k$). Observe that this function commutes with T. Now ask about the size of the set $\{n, f(n), f^2(n) \ldots\}$. This size function, too, commutes with T. Now we exploit the fact that $f(\mathfrak{m})$ is $T\mathfrak{m}$ to conclude that the size m of the set $\{\mathfrak{m}, f(\mathfrak{m}), f^2(\mathfrak{m})\ldots\}$ satifies $m = Tm + 1$, which is impossible.

This is actually a routine application of old ideas from the dawn of NF.

We need to recall the definition of separable subset. A is a separable subset of B iff $(\forall x \in B)(x \in A \vee x \notin A)$.

Let us write $\mathcal{SC}(X)$ for the set of separable subsets of X.

We'd better show that $\mathcal{SC}(X)$ is Nfinite if X is. $\mathcal{SC}(\emptyset)$ is $\{\emptyset\}$ which is Nfinite. (I know Michael states it above and i'm sure he's right, but i want to prove it myself for my own satisfaction.

Now sse X is Nfinite, $x \notin X$ and $\mathcal{SC}(X)$ is Nfinite. What about $\mathcal{SC}(X \cup \{x\})$? We need to check that a subset B of $X \cup \{x\}$ is separable iff there is a separable subset A of X s.t. $B = A \vee B = A \cup \{x\}$.

Suppose B is a separable subset of X. We desire that $B \cup \{x\}$ be a separable subset of $X \cup \{x\}$. So let y be an element of $X \cup \{x\}$. We want $y \in (B \cup \{x\}) \vee b \notin B \cup \{x\})$.

We have $y \in X \vee y = x$. $y = x$ gives $y \notin B$ so certainly $y \notin (B \cup \{x\})$ so that's OK.

If $y \in X$ then $Y \in B \vee y \notin B$ by separability of B. If $y \in B$ then certainly $y \in (B \cup \{x\})$ so that's OK. If $y \notin B$ then y isn't in $B \cup \{x\}$ either, beco's $y \neq x$.

For the other direction we desire that if Y is a separable subset of $X \cup \{x\}$ then $Y \cap X$ is a separable subset of X. Now if $z \in X$ we certainly have $z \in X \cup \{x\}$ so $z \in Y \vee z \notin Y$ by separability of Y. Since $z \in x$, $z \in Y$ implies $z \in (X \cap Y)$. If $z \notin Y$ then *a fortiori* $z \notin (X \cap Y)$.

■

I think it can't do any harm to prove that if B is a separable subset of X, and $x \notin X$, then B is a separable subset of $X \cup \{x\}$. But this is easy. If $z \in X \cup \{x\}$ then either $z \in X$ (in which case $z \in B \vee z \notin B$ by separability of B) or $z = x$ (in which case $z \notin B$) as desired.

So $\mathcal{SC}(X \cup \{x\})$ is the union of $\mathcal{SC}(X)$ and $\mathcal{SC}(X) \times \{x\}$, which are disjoint Nfinite sets, so it's Nfinite.

■

I think that last sentence tells us that if we take $2^{T|X|}$ to be $\mathcal{SC}(X)$ then $2^{n+1} = 2^n \cdot 2^n$

I've still got to get my head round separability. Are all subsets of singletons seperable subsets of that singleton? Sse $x \subset \{a\}$. Must we have $(\forall z \in \{a\}(z \in x \vee z \notin x)$? This is $(a \in x \vee a \notin x)$. No way José.

7.0.1 Two cute facts in strict iNF

We will show that no permutation moves the empty set and that no permutation has finite support.

First we record that

Lemma 24 *$\emptyset$ is the only x s.t. $(\forall a)(\neg\neg(a = x) \to a = x)$.*

Proof:

Clearly $\neg\neg(a = \emptyset)$ implies $\neg\neg(\forall x)\neg(x \in a)$ but we can import the $\neg\neg$ past the $\forall$ to obtain $(\forall x)\neg(x \in a)$ whence we obtain $a = \emptyset$ by extensionality.

For uniqueness we need to show that (unless double negation holds) $\emptyset$ is the only x s.t. $\neg\neg(y = x) \to y = x$.

To this end suppose A is nonempty and that $(\forall a)(\neg\neg(a = A) \to a = A)$. Assume $\neg\neg p$; we will deduce p. Consider $\{x \in A : p\}$. Since $\neg\neg p$, we have $\neg\neg(\{x \in A : p\} = A)$. This is beco's $p \to \{x \in A : p\} = A$ so, by contraposition, $\neg(\{x \in A : p\} = A) \to \neg p$. But $\neg\neg p$ by assumption so—contraposing again—we infer $\neg\neg(\{x \in A : p\} = A)$.

Now we use the assumption that $(\forall a)(\neg\neg(a = A) \to a = A)$ to infer $\{x \in A : p\} = A$ whence p. This justifies double negation. So, unless the law of double negation holds, A cannot be nonempty.

■

Lemma 25 *The support of a permutation is always a stable set: if π moves x then it must move everything notnotequal to x.*

Proof:

Suppose $\pi(a) \neq a$. Suppose $\not(a' = a)$ and $\pi(a') = a'$. Then $\pi(a) \neq a'$. Then if $a' = \pi(a')$, we have $\pi(a') \neq \pi(a)$ whence $a \neq a'$ contradicting the assumption that $\neg\neg(a = a')$.

So the support of π is a union of $(\neg\neg =)$-equivalence classes and is stable.

■

Theorem 26 *Unless double negation holds, $\emptyset$ is fixed by all permutations.*

Proof:

By lemma 25 we note that, for any permutation π at all, if $\neg\neg(a' = a)$ then $\neg\neg(\pi(a') = \pi(a))$.

Let π be a permutation. Suppose $\pi(a) = \emptyset$, and $\neg\neg(a' = a)$. Then $\neg\neg(\pi(a') = \pi(a))$. So $\neg\neg(\pi(a') = \emptyset)$. But by lemma 24 we have $\neg\neg(x = \emptyset) \to x = \emptyset$. So $\pi(a') = \emptyset$, and $a' = a$.

That is, the $(\neg\neg =)$-equivalence class of $\emptyset$ is $\{\emptyset\}$, and if $\pi(a) = \emptyset$ then the $(\neg\neg =)$-equivalence class of a is $\{a\}$. But by lemma 24 this implies $a = \emptyset$.

■

Theorem 27 *No finite set with at least two members can be stable.*

Proof:

First we have to show that there are no stable doubletons. To do that, we appeal to fishy sets. Consider the doubleton $\{a, b\}$, where $a \neq b$, and suppose it is stable. Let p be an arbitrary proposition. The set $\{x : x \in a \wedge p. \vee x \in b \wedge \neg p\}$

(call it 'f' for *fishy*) cannot be distinct from both a and b, so it is in the double complement of $\{a, b\}$. But $\{a, b\}$ is stable by assumption. So $f = a \vee f = b$, giving $p \vee \neg p$.

Then we do an induction.

Suppose X is not stable, and that $x \notin X$, but that $X \cup \{x\}$ *is* stable.

Suppose $\neg\neg(y \in X)$. Then certainly $\neg\neg(y \in X \cup \{x\})$, whence $y \in X \cup \{x\}$ by stability of $X \cup \{x\}$. Now $y \in \{x\}$ is impossible—since $\neg\neg(y \in X)$ but $x \notin X$—so we must have $y \in X$. Since y was arbitrary, we infer that X was stable after all, contradicting assumption. ■

This doesn't mean that there are no stable finite sets, $\emptyset$ is stable, and it allows that there might be stable singletons, and in fact there are: $\{\emptyset\}$ is stable. However—as we saw in lemma 24 above—it is the only one! It means only that there are no stable finite sets of size two upward.

Actually there is a much simpler proof. Any two members of a finite set are equal or unequal, so every finite set meets every $(\neg\neg =)$-equivalence class (if at all) on a singleton. And any stable set is a union of $(\neg\neg =)$-equivalence classes

Corollary 28 *(Unless double negation holds) there is no permutation of Nfinite support.*

If π moves a then it moves $\pi(a)$, so the support of π contains at least two things. Also, the support is stable. But no stable set with at least two members can be finite.

■

Can there be a permutation model containing a Quine atom if there is no Quine atom in the ground model? If so, then there is a permutation π and a set a s.t. π restricts to a bijection between the $(\neg\neg =)$-equivalence class of a and the $(\neg\neg =)$-equivalence class of $\{a\}$.

I think $\neg\neg(a' = a) \to \neg\neg)\{a'\} = \{a\})$ is safe. So ι injects the $(\neg\neg =)$-equivalence class of a into the $(\neg\neg =)$-equivalence class of $\{a\}$. But i don't think the injection is onto. Suppose

$$(\forall x)(\neg\neg(x = \{a\}) \to (\exists y)(\neg\neg(y = a) \wedge x = \{y\})$$

I don't believe this for a minute. Let X consist entirely of things notnotequal to a. Suppose X not equal to $\{a\}$. That is to say $\neg(\forall x)(x \in X \longleftrightarrow x = a)$. Bugger; that's not strong enough.

written up to clarify my thoughts and to amuse Randall.

Classically the situation *vis à vis* partitions and equivalence relations is clear. There is a 1-1 correspondence between partitions and equivalence relations. A partition (of V, for the moment) is a family of pairwise disjoint nonempty sets ("pieces") whose union is V. Thus Π is a partition iff $(\forall x)(\exists! X \in \Pi)(x \in X)$.

We say a partition Π_1 *is finer than* a partition Π_2 if every piece of Π_1 is a subset of a piece of Π_2. In these circumstances we also say Π_2 *is coarser than* Π_1. Notice that every partition maps onto any of its coarsenings.

Let us write '$\Pi(X)$' to denote the set of partitions of X. Let us write 'Π_a' to denote the partition corresponding to the equivalence relation $x \sim y$ iff $(x \in a \longleftrightarrow y \in a)$.

7.1 The Correspondence between Partitions and Equivalence Relations is constructively robust

Constructively the 1-1 correspondence holds up. Clearly if R is an equivalence relation then the set $\{[x]_R : x \in V\}$ is a partition within the meaning of the act. Conversely if Π is a partition then the relation $\{\langle x, y\rangle : (\forall p \in \Pi)(x \in p \longleftrightarrow y \in p)\}$ is an equivalence relation. It may be worth writing out (for the nervous) a proof that these two operations really are inverse (in this new constructive context) in the way one expects. Let's do it ...

THEOREM 29 *Constructively the operations taking an equivalence relation to a partition and vice versa are mutually inverse.*

Proof:

Let's start with an equivalence relation R, and go to the partition and back—and hope we end up where we started. The corresponding partition is $\Pi = \{[x]_R : x \in V\}$. Let's extract an equivalence relation from it. We will say that x is related to y iff they belong to the same pieces of Π, which is to say

$$(\forall z)(x \in [z]_R \longleftrightarrow y \in [z]_R)$$

which in turn is equivalent to

$$(\forall z)(R(x, z) \longleftrightarrow R(z, y))$$

Now substituting x/z gives

$$R(x, x) \longleftrightarrow R(x, y)$$

which of course is just $R(x, y)$.

Starting with a partition Π we obtain the equivalence relation $R(x, y) \longleftrightarrow (\forall p \in \Pi)(x \in p \longleftrightarrow y \in p)$. What is a piece of this partition? Suppose $x \in p \in \Pi$; what is $[x]$? It is $\{y : (\forall p \in \Pi)(x \in p \longleftrightarrow y \in p)\}$.

We will show that this object is p by proving two inclusions. Suppose $y \in p$. Then we have x and x both in p, so certainly $(\forall p \in \Pi)(x \in p \longleftrightarrow y \in p)$ which puts y into $[x]$.

For the other inclusion suppose $y \in [x]$; but then $y \in p$ is immediate.

■

But that's the end of the good news. Classically any two pieces of a partition are disjoint or identical. Constructively we can't prove this, tho' we can prove that if they meet they are identical. We're going to need an adjective for partitions that have this classical property; i'm going to call them *hard* (short for "hard-edged").

The assertion that the partition ι"V is hard is just *Tertium Non Datur* for $=$, and is equivalent to classical logic. So we have to be careful!

The double complement of an equivalence relation is an equivalence relation. This is a simple corollary of lemma 4.

In contrast

REMARK 30

(i) If Π and $\sim\sim\Pi$ are both partitions then they are equal, and
(ii) that can happen only if we have excluded middle.

Proof:

(i) Suppose Π and $\sim\sim\Pi$ are both partitions. Let x be arbitrary. It must belong to some piece X of $\sim\sim\Pi$. It must also belong to a piece X' of Π. Now $\Pi \subseteq \sim\sim\Pi$ so $X' \in \sim\sim\Pi$. Now $\sim\sim\Pi$ is a partition, so we must have $X = X'$. So every piece in $\sim\sim\Pi$ also belongs to Π, so they're identical!

(ii) Suppose $\Pi = \sim\sim\Pi$, and that p_1 and p_2 are two pieces. Consider the fishy set $\{x : ((x \in p_1) \wedge q) \vee ((x \in p_2) \wedge \neg q)\}$; since $\Pi = \sim\sim\Pi$ this fishy set is a piece of Π, by lemma 10. It is inhabited, by a say. Then $((a \in p_1) \wedge q) \vee ((a \in p_2) \wedge \neg q)$. This enforces $q \vee \neg q$.

■

We know that partitions and equivalence relations come in pairs. What sort of pairs can there be?

The pair might have no nontrivial properties at all. There seems to be three grades of niceness:

DEFINITION 31

(A) *The partition is* hard*: any two pieces are equal or disjoint. The corresponding equivalence relation satisfies $(\forall x, y)(R(x, y) \vee \neg R(x, y))$. This is of course the same as the partition being* discrete*: any two inhabitants are equal or unequal.*

(B) *Any two notnotequal pieces of the partition are equal; If $\neg\neg R(x, y)$ then $R(x, y)$; $R = \sim\sim R$ (R considered as a set of ordered pairs).*

(C) *The partition is a coarsening of $V/(\neg\neg =)$; in other words $\neg\neg(x = y) \to R(x, y)$.*

LEMMA 32

We will show that all the conditions in (A) are equivalent, (B) similarly. It's pretty clear that (A) implies (B)[5]*. and that (B) implies (C).*

[5] I used to think that (B) and (C) were equivalent but they aren't, beco's $\sim\sim$does not commute with $\bigcup$: $\sim\sim\bigcup_{i \in I} A_i$ is not reliably the same as $\bigcup_{i \in I} \sim\sim A_i$.

Proof:

(A) is equivalent to both

(i) ("any two pieces are disjoint or identical") and

(ii) ("any two things are equivalent or inequivalent"):

(i) implies (ii). Think of a point x and a piece p. Then ask whether or not $p = [x]$. If $p = [x]$ then $x \in p$; if $p \neq [x]$ then $x \notin p$.

(ii) implies (i). Given p_1 and p_2 pick $x \in p_1$. By (ii), $x \in p_2 \vee x \notin p_2$. One horn gives $p_1 = p_2$; the other gives $p_2 \neq p_1$.

(B) is equivalent both to

(iii) $(\forall p \in \Pi)(\,\sim\sim p = p)$; and to

(iv) $(\forall p_1 p_2 \in \Pi)(\neg\neg(p_1 = p_2) \to p_1 = p_2)$.

(iii) implies (iv). Suppose Let p and q be two pieces, with $\neg\neg(p = q)$. Anything notnot in one is notnot in the other, but anything notnot in one is actually *in* the one. So they are identical by extensionality.

(iv) implies (iii). Suppose $x \in \sim\sim p$. Then we cannot have $[x]_R \neq p$. So $\neg\neg([x]_R = p)$, whence $[x]_R = p$ by (iv), giving $\sim\sim p = p$.

First we note that (A) implies (B), specifically that (ii) implies (iii).

Suppose Π satisfies (ii), and let p be a piece. Consider $x \in \sim\sim p$. Now $[x]$ and p are either disjoint or equal. They cannot be disjoint beco's x is notnot in the intersection. So they are identical.

(iv) implies (C) as follows:

If $\neg\neg(x' = x)$ then $\neg\neg R(x, x')$ using substitution on $R(x, x)$. But if $\neg\neg R(x, y) \to R(x, y)$ always then we infer $R(x, x')$. So (iv) implies that the partition is a coarsening of $V/(\neg\neg =)$. ■

Should Say something about whether (A)-flavoured partitions are coarser or finer than (B)-flavoured partitions. My head is spinning!

There are some partitions discussion of which might help to concentrate the mind.

(i) ι"V. This is not nice in any of these three senses.

(ii) $V/(\neg\neg =)$. This is nice in sense (C) obviously(!) but also in sense (B). It won't be (A)-nice unless we have $(\forall xy)(\neg(x = y) \vee \neg\neg(x = y))$.

(iii) there is also the partition pertaining to the equivalence relation $x \sim y$ iff $(\forall z)(z \notin x \longleftrightarrow z \notin y)$. This appears to be a proper coarsening of $\neg\neg =$ but we need to check.

(iv) Consider the pair $\{\{\emptyset\}, V \setminus \{\emptyset\}\}$. Is it a partition? One would need $(\forall x)(x = \emptyset \vee x \neq \emptyset)$, and there doesn't seem to be any hope of that.

(v) For each set a there is the relation $R(x, y) \longleftrightarrow (x \in a \longleftrightarrow y \in a)$. In the case where $a = \{\emptyset\}$ the quotient (which we are notating Π_a) is precisely Ω, which doesn't look particularly (B)-like.

Worth observing that Π_X is the coarsest partition containing X as a piece. Indeed we can prove something more general:

Remark 33 *If $A \subseteq X$ then the partition of X corresponding to $x \sim y \longleftrightarrow (x \in A \longleftrightarrow y \in A))$ is the coarsest partition of X of which A is a piece.*

Proof: Let Π be a partition of X of which A is a piece. If x and y belong to the same piece of Π we are saying that $(\forall p \in \Pi)(x \in p \longleftrightarrow y \in p)$. In particular, since A is a piece of Π, we have $x \in A \longleftrightarrow y \in A$, so x and y belong to the same piece of the partition of X corresponding to $x \sim y \longleftrightarrow (x \in A \longleftrightarrow y \in A))$. ■

I think the moral of this is that basically partitions are never kfinite unless the logic is classical. Is there any set (other than V) with a genuine complement? i can't find one ($\{\emptyset\}$ was a feeble attempt) but nor can i think of any reason why there should not be one!

Well, suppose $(\forall x)(x \in a \vee x \notin a)$. Consider the equivalence relation $x \sim y \longleftrightarrow (x \in a \longleftrightarrow y \in a)$ ... yes it has precisely two pieces. But is there such an a?

Π_a must contain both a and $V \setminus a$. Its sumset must be V so, unless $(\forall w)(w \in a \vee w \notin a)$, it must have some some extra members. But these extra members cannot contrive to be distinct both from a and from $V \setminus a$. Suppose $x \in p$, with $p \neq a$ and $p \neq V \setminus a$. Then $x \notin a$ but also $x \notin V \setminus a$ when $\neg\neg(x \in a)$, a contradiction.

Remark 34 *Every refinement of $V/(\neg\neg =)$ maps onto Ω.*

Proof:

If Π is a partition that refines $V/\neg\neg =$ then whenever x, y belong to $p \in \Pi$, we have $\neg\neg(x = y)$. But if $\neg\neg(x = y)$ then $x = \emptyset$ iff $y = \emptyset$. This is beco's $\neg\neg(x = \emptyset)$ implies $x = \emptyset$. ($\neg\neg(x = \emptyset)$ is $\neg\neg(\forall y)(y \notin x)$ and this implies $(\forall y)(y \notin x)$.) So if $p \in \Pi$ then $[[x = \emptyset]]$ is the same for all $x \in p$. Recall that $[[\phi]]$ (the truth-value of ϕ) is $\{x : (x = \emptyset) \wedge \phi\}$, so that $[[\phi]]$ is always a subset of $\{\emptyset\}$ and is a member of $\mathcal{P}(\{\emptyset\}) = \Omega$.

So, send each $p \in \Pi$ to the truth-value $[[x = \emptyset]]$ for any (all) $x \in p$. This maps Π onto Ω.

Just need to check that it really is onto .... If $v \in \Omega$ then v is the destination of $[v]$, as follows. $[v]$ gets sent to $[[v = \{\emptyset\}]]$ and that is $\{y : y = \emptyset \wedge (v = \{\emptyset\})\}$ which should be just v.

We desire:

$$\{y : y = \emptyset \wedge v = \{\emptyset\}\} = v.$$

This holds iff

$$(\forall y)((y = \emptyset \wedge v = \{\emptyset\}) \longleftrightarrow y \in v)$$

which we prove as follows:

R to L:

If $y \in v$ then $y = \emptyset$ (beco's $v \subseteq \{\emptyset\}$) and so $v = \{\emptyset\}$;

L to R is easy.

■

Actually this is overkill: every refinement of $V/(\neg\neg =)$ maps onto $V/(\neg\neg =)$, so all we have to do is show that $V/(\neg\neg =)$ maps onto Ω.

For any set x we can consider the pair $\{x, V \setminus x\}$, but this is not a partition unless $(\forall y)(y \in x \vee y \notin x)$. Come to think of it, are there any partitions of V into two pieces? (No piece of a partition is allowed to be empty!) Revisit this in connection with Π_a.

Can one show that no finite partition has decidable equality between its pieces?

Suppose $p \in \Pi$, Π a partition. Can we prove $(\forall x)(x \in p \vee x \in \bigcup(\Pi \setminus \{p\}))$? Every piece of a partition is a piece of a two-piece partition? Surely not.

Here's something with some bite. For any a consider the equivalence relation $x \in a \longleftrightarrow y \in a$; let's call it $\simeq_a$. Consider the quotient over $\simeq_a$; let's call it Π_a. I think the partition corresponding to $\sim\sim(\simeq_a)$ is $\Pi_{\sim\sim a}$. Perhaps this notation is not the most felicitous... (!)

There is a map from Π_a to Ω given by $x \mapsto [[x \in a]]$. Reflect that if x and y belong to the same piece of Π_a then they get sent to the same element of Ω, so this map *factors through* (i think that's the phrase) the quotient map. So there is an injective map $\Pi_a \to \Omega$; but is it onto? Given a set a and a truth-value $p \in \Omega$ we seek x s.t. $[[x \in a]] = p$. In full this is $p = \{y : y = \emptyset \wedge x \in a\}$, and there's no obvious reason why there should be such an x. Indeed, if a is something hard like $\{\emptyset\}$ then there is almost certainly no surjection $\Pi_a \twoheadrightarrow \Omega$. Sort this out and draw a picture

Remark 35 *Π_a is the coarsest partition of which a is a piece.*

Proof:

Let Π be a partition and a a piece of Π. Then there is a surjection $s : \Pi \twoheadrightarrow \Pi_a$ as follows. Suppose x and y belong to the same piece p of Π. That is to say that for all $b \in \Pi$, $x \in b \longleftrightarrow y \in b$ so in particular (since a is a piece of Π) we have $x \in a \longleftrightarrow y \in a$ so y and y belong to the same piece of Π_a. Call this piece $s(p)$.

■

Thinking about Π_a may resolve the question of whether or not the set of partitions of a kfinite set is kfinite (or perhaps whether or not the set of partitions of an Nfinite set is Nfinite). If every subset of X can be a piece of a partition of X then $\Pi(X)$ cannot be relied upon to be kfinite (let alone Nfinite) if X is, beco's $\bigcup\Pi(X)$ would be a union of a kfinite set of kfinite sets and would therefore be kfinite. It does seem clear that, for any $a \subseteq X$, Π_a is a partition of which a is a piece, so every subset of X is a piece of some partition of X. So the question becomes, exactly where does the proof go wrong?

7.2 Partitions of Kuratowski-finite sets

The set of partitions of a Kfinite set; is it kfinite? What about Nfinite sets?

Several facts to bear in mind

• Every partition of a kfinite set is kfinite, being a quotient of a kfinite set. Quotients of Nfinite sets are not reliably Nfinite.

• The set of partitions of $X \cup \{x\}$ is a quotient of $(X \cup \{x\})\times$ the set of partitions of X, so it looks as if we have the makings of a proof by induction that the set of partitions of a kfinite (Nfinite?) set is kfinite (Nfinite?)

but

• The sumset of the set of partitions of a kfinite set would be kfinite, and seems to suggest that the power set of a kfinite set is kfinite. We would then have some explaining to do, beco's we have counterexamples to that. Perhaps not every subset of X is a piece of a partition of X? But, if $Y \subseteq X$, $y_1 \sim y_2$ iff $y_1 \in Y \longleftrightarrow y_2 \in Y$ is an equivalence relation and Y is an equivalence class. We show in section 7.1 that the correspondence between partitions and equivalence relations is constructively robust.

Suppose X and $\Pi(X)$ are both kfinite. Then $\Pi(X) \times X$ is also kfinite, being the product of two kfinite sets. We want $\Pi(X \cup \{x\})$ to be kfinite. An element p of $\Pi(X) \times X$ is a partition of X paired with an element of X, and that member of X identifies a piece of that partition, so p can be thought of as a partition of X with a designated element. Now consider the function that takes that decorated partition and inserts the new element x into the designated element. The image of $\Pi(X) \times X$ in this function is the set of those partitions of $X \cup \{x\}$ where the piece containing x is not a singleton, so this image does not include all partitions. However it is kfinite. There is also the kfinite set $\Pi(X) \times \{x\}$, and it is in 1-1 correspondence with the set of partitions of $X \cup \{x\}$ where the piece containing x is a singleton.

[later] But isn't $\Pi(X \cup \{x\})$ the union of these two sets? And aren't they disjoint? And isn't the union of tow disjoint kfinite sets kfinite? But we need to know whether or not $x \in X$.

But perhaps the set of kfinite partitions of a kfinite set is Kfinite, or the set of partitions into kfinite pieces. Somethiong along those lines ought to be true?

7.3 Permutations and Partitions

Definition 36 *A permutation [well, at least a permutation that is an involution] is a set π of pairs such that*

$$(\forall x)((\exists! p_1 \in \pi)(x = \mathtt{fst}(p_1)) \wedge (\exists! p_2 \in \pi)(x = \mathtt{snd}(p_2)))$$

$\mathbb{1}$ *is the identity permutation of V.*

Total functions $f : V \to V$ are good sources of both partitions and permutations.

Given such an f, consider E_f the $\subseteq$-least set containing V and closed under $X \mapsto f``X$. Then we have the equivalence relation $x \sim_f y$ iff $(\forall X \in E_f)(x \in X \longleftrightarrow y \in X)$. What can we say about this partition? Observe that $f \restriction \bigcap E_f$ is a permutation of $\bigcap E_f$.

The appearance of the uniqueness quantifiers should alert the reader to the thought that permutations resemble partitions rather than equivalence relations. The proof of the following remark is very like the proof of remark 30.

REMARK 37 *The double complement of a permutation is never a permutation unless the logic is classical.*

Proof:

Suppose τ and $\sim\sim\tau$ are both permutations.

Suppose $\neg\neg(y' = y)$. Then, for some x, $\langle x, y\rangle \in \tau$. We have $\neg\neg(\langle x, y\rangle = \langle x, y'\rangle)$, so $\neg\neg(\langle x, y'\rangle \in \tau)$, which is to say $\langle x, y'\rangle \in \sim\sim\tau$. But $\langle x, y\rangle \in \tau \subseteq \sim\sim\tau$. Now $\sim\sim\tau$ is a permutation by assumption so we must have $y = y'$. But y was arbitrary. So $=$ is a stable relation, which is a form of classical logic.

■

Observe that we didn't assume that the permutation was nontrivial. The identity relation is a permutation, and its double complement is of course $\neg\neg =$ which is not a permutation.

The connection between partitions and permutations is that

(i) the cycles of a permutation form a partition of V.

(ii) Every group of permutations of V partitions V into orbits.

It's this connection to partitions (which are constructively problematic as we have seen) that makes me wonder whether *i*NF proves that there are any nontrivial permutations at all!

The set of permutations that are notnot equal to $\mathbb{1}$ is presumably $\sim\sim\{\mathbb{1}\}$ and is a normal subgroup. Let us call this group $\mathcal{I}$.

$(\forall x, y)(\neg\neg(x = y) \to \neg\neg(f(x) = f(y)))$ so every permutation in $\mathcal{I}$ can be tho'rt of as acting on the $\neg\neg$ =-equivalence classes, so we might as well restrict our attention to the quotient.

Elements of $\mathcal{I}$ are of no use from the point of view of consistency proofs; they are so like $\mathbb{1}$ that they won't change anything.

The question remains: is $\mathcal{I}$ the whole group? Consider the function $f \mapsto \{x : x = 0 \wedge f = \mathbb{1}\}$. This is a map from $\mathcal{I}$ to Ω. Does it show that $\mathcal{I}$ is not kfinite?

In *i*NF we presumably cannot show that $\mathrm{Symm}(V)$ has precisely one orbit. Presumably assertions that the set of orbits is in some sense small will have logical force.

It would be nice if $\neg\neg(\mathcal{I} = \{\mathbb{1}\})$ but of course there's no reason to expect that. If that were true we could argue as follows

$(\forall xy)(x$ and y belong to the same $\{\mathbb{1}\}$ orbit iff $x = y)$.

But $\neg\neg\{\mathbb{1}\} = \mathcal{I}$ whence

$\neg\neg(\forall xy)(x$ and y belong to the same $\mathcal{I}$ orbit iff $x = y)$;

and (import $\neg\neg$)

$(\forall xy)\neg\neg(x$ and y belong to the same $\mathcal{I}$ orbit iff $x = y)$;

which (i think) would give

$(\forall xy)(x$ and y belong to the same $\mathcal{I}$ orbit iff $\neg\neg(x = y))$

which says that $\mathcal{I}$-orbits are just the $\neg\neg$ =-equivalence classes.

The relation "Every set closed under both f and f^{-1} containing either x or y contains the other" is an equivalence relation. But is it a *stable* equivalence relation?

$\mathrm{Symm}(V)/\mathcal{I}$ acts on the $\neg\neg$ =-equivalence classes.

Now sse $\neg\neg(\tau = \mathbb{1})$. All τ can do is move things around within $\neg\neg$ =-equivalence classes. So every τ-cycle is a subset of a $\neg\neg$ =-equivalence class. Can it be a *proper* subset? Sounds unlikely ... how can τ distinguish things that are notnotequal?

Suppose τ has two orbits o_1 and o_2 that are included in the same $\neg\neg$ =-class. Every member of one is notnot= to every member of the other. So they notnot meet. So they are notnot equal. That seems to be the best we can do.

Suppose f is a nontrivial permutation but that there is no interpretation of Heyting arithmetic, and consider an arbitrary x. Does x belong to the f-closure of $\{f(x)\}$? If it doesn't, then we have interpretation of Heyting arithmetic. So the conclusion must be that $\neg\neg(x \in$ the f-closure of $\{f(x)\})$.

If we have a permutation that $\neg\neg$ = the identity then all its cycles are subsets of $\neg\neg$ =-classes. Such a permutation seems to contain information that discriminates among things $\neg\neg =$ each other: $f(x)$ is *one* of the things $\neg\neg = x$.

Suppose π is a permutation that maps A onto B. What does it do to $\sim\sim A$? We have

$a \in A \to (\exists b \in B)(\lambda(\pi, a, b) \wedge (\forall b')(\lambda(\pi, a, b') \to b = b'))$

... where $\lambda(\pi, a, b)$ says $\pi(a) = b$

Now suppose $a' \in \sim\sim A$. Then there is a unique b' with $\lambda(\pi, a', b')$. Suppose this $b' \notin B$. Then it cannot be π of anything in A: $\pi^{-1}(b') \notin A$, contradicting $\pi^{-1}(b') = a' \in \sim\sim A$. So $b' \in \sim\sim B$. So $\pi : \sim\sim A \twoheadrightarrow \sim\sim B$.

I don't suppose there is a converse but it could be interesting if there were. The first point to make is that one cannot recover A from $\sim\sim A$. However if c is a fishy combination of a and b then $f(c)$ is a fishy combination of $f(a)$ and $f(b)$

Is it true that every member of $\sim\sim X$ is a fishy combination of two members of X? $\{x : (x \in a \wedge p) \vee (x \in b \wedge \neg p)\}$. We can express this without the letter 'p' by replacing 'p' by '$s = \{\emptyset\}$' thus:

$(\forall y \in \sim\sim X)(\exists s \subseteq \{\emptyset\})(\exists a, b \in X)(\forall w)(w \in y \longleftrightarrow (w \in a \wedge s = \{\emptyset. \vee w \in b \wedge s \neq \{\emptyset\})$

We're considering orbits of $\mathcal{I}$. The orbit of x is a subset of $[x]_{\neg\neg=}$. Indeed $[x]_{\neg\neg=}$ is partitioned into orbits. The orbits are all notnotequal. After all, if i am equal to you, then my orbit is equal to your orbit; so if i am notnotequal to you, then my orbit is notnotequal to your orbit. So: for all orbits xy, $\neg\neg(x = y)$. But we can't pull the notnot to the front, so for all we know it might be the case that $\neg(\forall \text{ orbits } xy)(x = y)$. Now, if the quotient (the set of orbits) is Nfinite then any two things in it are identical. Kfinite implies notnotfinite so if the quotient is kfinite then notnot any two things in it are identical.

I think this is what is going on. Suppose f is a stable permutation, and y is any set. There there is x such that $f(x) = y$. Now suppose y' satisfies $\neg\neg(y = y')$; then we must have $f(x) = y'$ by substitutivity of $\neg\neg =$ (since f is stable). Now since f is a function we must have $y = y'$. So what does this prove? I think it proves that if there is a stable permutation then excluded middle holds. If that sounds a bit much just reflect that equality is not stable!

Some things are notnotequal only to themselves.
Anything notnot= to $\emptyset$ is empty: $\neg\neg(x = \emptyset)$ implies
$\neg\neg(\forall y)(y \in x \to y \in \emptyset)$ implies
$(\forall y)\neg\neg(y \in x \to y \in \emptyset)$ implies
$(\forall y)(\neg\neg(y \in x \to \neg\neg(y \in \emptyset)))$ implies
$(\forall y)(\neg\neg(y \in x \to \perp))$ implies
$x = \emptyset$

How about $\neg\neg(x = \{\emptyset\})$?
$\neg\neg(\forall y)(y \in x \longleftrightarrow y = \emptyset)$
$(\forall y)\neg\neg(y \in x \longleftrightarrow y = \emptyset\})$
$(\forall y)(\neg\neg(y \in x) \longleftrightarrow \neg\neg(y = \emptyset))$
$(\forall y)(\neg\neg(y \in x) \longleftrightarrow y = \emptyset)$
which clearly implies $x = \{\emptyset\}$.

7.4 Rieger-Bernays methods and suchlike

I think Boffa's lemma works. "Dense Nfinite" is a 1-formula, so if σ is a permutation and x is a dense Nfinite set then $\sigma``x$ is dense Nfinite too ... and therefore notnotequal to x, one might add. The thought was that this might tell us how many dense Nfinite sets there but it doesn't seem to help.

7.5 A Conversation with André 28/iv/19

I am trying to convince him that there are hardly any permutations in the *i*NF world. Simplest case: $a \neq b$. How about the permutation

$$\{\langle x, y\rangle : (x = a \wedge y = b) \vee (x = b \wedge y = a) \vee x = y\}?$$

We need to show that this defines a map which is single valued and surjective. I'm not going to worry about the surjectivity co's the symmetry of the formula should see to that. But is it functional? Suppose for a given x we can find y and y' such that

$(x = a \wedge y = b) \vee (x = b \wedge y = a) \vee x = y$ and

$(x = a \wedge y' = b) \vee (x = b \wedge y' = a) \vee x = y'$

Then it's easy to show that $y = y'$. So it's functional. But i don't think we can show that there reliably is such a y (or y') in the first place.

So the thought is that one should attempt to prove

$$a \neq b \wedge (\forall x)(\exists! y)[(x = a \wedge y = b) \vee (x = b \wedge y = a) \vee x = y]. \rightarrow .(\forall z)(z = a \vee z = b \vee (z \neq a \wedge z \neq b))$$

Assume the antecedent, and pick a random z with a view to doing a UG. By uniqueness there is a unique such y which we can write '$f(z)$'. The disjunction gives us

$(z = a \wedge f(z) = b) \vee (z = b \wedge f(z) = a) \vee z = f(z)$.

The first disjunct gives us $z = a$, the second gives us $z = b$ and the third gives $z \neq a \wedge z \neq b$.

So, yes, i was right: there is a permutation swapping a and b and fixing everything else only if a and b are isolated.

There is still the challenge of showing that the existence of a nonidentity permutation has nontrivial logical consequences.

If Π is a partition of X is $\sim\sim\Pi$ a partition of $\sim\sim X$? I bet it isn't ...

8 The Truth-Value Algebra

We start off with a general observation about relations between constructive and classical theories.

Theorem 38 *For any set theory T in which Ω is a set, adding the principle*

$$\forall\neg\neg \;\rightarrow\; \neg\neg\forall$$

gives a system as strong as that version of T that allows excluded middle for those formulæ for which it admits comprehension.

Proof:

We know constructively that there are not three distinct truth-values: (the sequent $\neg(A \longleftrightarrow B), \neg(B \longleftrightarrow C), \neg(C \longleftrightarrow A) \vdash$ is intuitionistically valid) so we can prove this fact in T obtaining

$$T \vdash (\forall x \in \Omega)\neg(x \neq \bot \wedge x \neq \top).$$

This is constructively the same as

$$T \vdash (\forall x \in \Omega)(\neg\neg(x = \bot \vee x = \top))$$

which is

$$T \vdash (\forall x)(x \in \Omega \to \neg\neg(x = \bot \vee x = \top))$$

which implies (since constructively we have $(A \to \neg\neg B) \to \neg\neg(A \to B)$)

$$T \vdash (\forall x)\neg\neg(x \in \Omega \to (x = \bot \vee x = \top)).$$

Adding commutation-of-$\forall$-with-$\neg\neg$ to T would now give us

$$\neg\neg(\forall x)(x \in \Omega \to (x = \bot \vee x = \top))$$

which is of course the same as

$$\neg\neg(\forall x \in \Omega)(x = \bot \vee x = \top)$$

and we can consistently add

$$(\forall x \in \Omega)(x = \bot \vee x = \top)$$

to T to obtain a theory which we can call T^*.

Therefore, if ϕ is a formula s.t. T proves $\{x : \phi\}$ exists, then T^* proves $\phi \vee \neg\phi$. ■

The point is that although adding commutation-of-$\neg\neg$-with-$\forall$ to constructive predicate logic does not give classical predicate logic, it does give classical logic in the presence of set-theoretic axioms that enable us to reason about truth-values as objects of the theory.

I have been careful not to say that it gives a theory as strong as the classical version of the theory T that we started with. If T is a constructive version of a set theory with a separation scheme then this does happen. However in the case of interest here—which is of course *i*NF and NF—all it gives is the relative consistency of *i*NF + excluded middle for weakly stratified formulæ. (see remark 64). In fact commutation does enable us to give an interpretation of full NF, but this is for other, rather special, reasons. (see section 11).

Finally let us note that

REMARK 39
Commutation-of-$\forall$-with-$\neg\neg$ (for stratified formulæ) is equivalent to the principle

$$(\forall x)(\neg\neg(x = \sim\sim x))$$

Proof: Assume $(\forall x)(\neg\neg F(x))$. This is $\{x : \neg\neg F(x)\} = V$. Now $\{x : \neg\neg F(x)\} = \sim\sim\{x : F(x)\}$. By the commutation principle we infer $\neg\neg(\sim\sim\{x : F(x)\} = \{x : F(x)\})$. So $\neg\neg(V = \{x : F(x)\})$, which is to say $\neg\neg(\forall x)(F(x))$. The other direction is easy. ■

The truth-value algebra is strongly cantorian

In NF, sets x such that the restriction of the singleton function to x exists are said to be *strongly cantorian*. The following observation seemed very striking at the time, but there has been no fall-out from it.

Remark 40 Ω *is strongly cantorian.*

Proof:

Ω is $\mathcal{P}$ of a singleton; singletons are strongly cantorian and power sets of strongly cantorian sets are strongly cantorian, even constructively. ■

Classically, strongly cantorian sets are small, so this appears to be telling us that there are not very many truth-values. If there are few truth-values one starts to think that the logic is classical.

9 Finite Sets

Stuff to fit in

Things i learnt from Michael B:

(i) The importance of separable subsets

(ii) The possibility of making use of Church numerals.

Neither of them have come to anything—yet. But they yet might.

(A result of trying to explain to Michael what has gone wrong with his doomed proof)

Michael is interested in $\mathfrak{m}$, the last Nfinite cardinal.

We define $0 = \{\emptyset\}$; $S(x) = \{Y \cup \{y\} : y \notin Y \in x\}$. We prove by induction that all successors are Nfinite cardinals. If you are a member of the last cardinal then your double complement is V. So all the things in $\mathfrak{m}$ have double complement=V, they are all bijective copies of one another. And anything not-notequal to a member of $\mathfrak{m}$, or the same size as a member of $\mathfrak{m}$ is a member of $\mathfrak{m}$? Yes, i think so. (Are they all 1-equivalent?)

Now consider not the successors, but the successors quotiented out by $\neg\neg =$. 0 and $\mathfrak{m}$ now both have only one element. It looks like a row of Pascal's triangle. $\forall x \in n)(\exists m \in \mathfrak{m}-n)(\sim\sim m = V \setminus x)$. Using this machinery we can say whether $\mathfrak{m}$ is even or odd. I think we can show that every successor is even or odd.

We need to think about selection functions of $V/\neg\neg =$

in NF(U) (or INF) the church numerals correspond to closed stratified set abstracts, as do church successor, addition and multiplication. Beeson is trying to persuade me that this will give a smooth proof of the axiom of infinity in INF. I am intruiged but unhappy. Since infinity is not a theorem of NFU we know that this cannot work in NFU. Presumably the obstacle lies in showing that succ is injective. I think i can see how to prove in NF that succ is injective, but basically that involves throwing away the church numerals. Let me explain. If injectivity of succ fails then there is some n s.t for all f, $f^n$is idempotent. But if there are infinitely many equipollence classes of inductively finite sets (as we can prove there to be in NF) then the successor function on these equipollence classes is a counterexample. So succ is injective after all. This is OK, but it works by going back to Frege number, and the point of the Church numeral ruse was to evade that. Does this mean that Church numerals are a red herring?

In a negative interpretation one of course needs an equivalence relation to serve as equality in the target language. One obvious one is $\neg\neg(x = y)$; another is $\sim\sim x = \sim\sim y$. It would be nice if they were the same but of course they aren't—unless we have "commutation": $\forall\neg\neg \rightarrow \neg\neg\forall$. So it's not surprising that commutation gives us full NF. A helpful (scene-setting) triviality is the following.... Let us recall the familiar definition of $\sim^+$:

$$X \sim^+ y \longleftrightarrow [(\forall x \in X)(\exists y \in Y)(x \sim y) \wedge (\forall y \in Y)(\exists x \in X)(x \sim y)]$$

The cute fact is that the equivalence relation $\sim\sim x = \sim\sim y$ is + of the equivalence relation $\neg\neg(x = y)$. (Or will be when we give it a constructive tweak)

Suppose

$$(\forall x \in X)(\exists y \in Y)(\neg\neg(x = y)) \wedge (\forall y \in Y)(\exists x \in X)(\neg\neg(x = y)))$$

we want $\sim\sim X = \sim\sim Y$.

Suppose $\neg\neg(x \in X)$. $x \in X$ implies $(\exists y \in Y)(\neg\neg(x = y)$, so $\neg\neg(x \in X)$ implies $\neg\neg(\exists y \in Y)(\neg\neg(x = y)$. Now $\neg\neg\exists\neg\neg$ implies $\neg\neg\exists$ so we infer $\neg\neg(\exists y \in Y)(x = y)$. Now $(\exists y \in Y)(x = y)$ implies $x \in y$ so $\neg\neg(\exists y \in Y)(x = y)$ implies $\neg\neg(x \in y)$.

Hmmm... That was only one direction. The other might not work...

Assume $\sim\sim X = \sim\sim Y$. We want

$$(\forall x \in X)(\exists y \in Y)(\neg\neg(x = y)) \wedge (\forall y \in Y)(\exists x \in X)(\neg\neg(x = y)))$$

It's not going to work. Perhaps we want to define $\sim^+$ instead as follows

$$X \sim^+ y \longleftrightarrow [(\forall x \in X)\neg\neg(\exists y \in Y)(x \sim y) \wedge (\forall y \in Y)\neg\neg(\exists x \in X)(x \sim y)]$$

Yes, *that* works!

Perhaps $\sim\sim X = \sim\sim Y$ is $(\forall x \in X)\neg\neg(\exists y \in Y)(\neg\neg(x = y)) \wedge (\forall y \in Y)\neg\neg(\exists x \in X)(\neg\neg(x = y))$

Sse $\neg\neg(x \in X)$. Then $\neg\neg(\exists y \in Y)(\neg\neg(x = y))$ and $\neg\neg(\exists y \in Y)(x = y)$ whence $\neg\neg(x \in Y)$.

Yep

Remark 43 says that if A and B are both Nfinite sets whose double complements are both V then they are notnotequal. I think we can prove something much stronger than that.

Let us write '$x =_0 y$' for $\neg\neg(x = y)$ and thereafter write $x =_{n+1} y$ for $x(=_n)^+y$ …*i.e.* $x =_{n+1} y$ iff $\neg\neg(\forall x' \in x)(\exists y' \in y)(x' =_n y') \wedge (\forall y' \in y)(\exists x' \in x)(x' =_n y')$. This is a metalanguage recursion over all concrete naturals.

(I do not know how to get rid of the '$\neg\neg$'.

I claim:

THEOREM 41 *If A and B are both Nfinite sets whose double complements are both V then $x =_n y$ for all concrete n.*

Proof:

We start by observing that $=_n$ is reflexive.

Suppose A and B are both Nfinite sets whose complement is V.

$A =_{n+1} B$ is

$(\forall x \in A)(\exists y \in B)(x =_n y) \wedge (\forall y \in B)(\exists x \in A)(x =_n y)$

First we tackle the left conjunct, and prove $(\forall x \in A)\neg\neg(\exists y \in B)(x =_n y)$

Let $x \in A$ be arbitrary. Suppose $\neg(\exists y \in B)(x =_n y)$. This contradicts $x \in B$, since x would be a witness to the $\exists y$ ($=_n$ is reflexive after all), giving $x \notin B$. But $\neg\neg(x \in B)$ by assumption, whence $\neg\neg(\exists y \in B)(x =_n y)$ and UG gives $(\forall x \in A)\neg\neg(\exists y \in B)(x =_n y)$ as desired. But A is Nfinite so we can use Linton-Johnstone to infer $\neg\neg(\forall x \in A)\neg\neg(\exists y \in B)(x =_n y)$.

We treat the other conjunct similarly, obtaining $\neg\neg(\forall y \in B)\neg\neg(\exists x \in A)(x =_n y)$. Putting the two together we obtain

$\neg\neg(\forall x \in A)\neg\neg(\exists y \in B)(x =_n y) \wedge \neg\neg(\forall y \in B)\neg\neg(\exists x \in A)(x =_n y)$, which is to say

$\neg\neg((\forall x \in A)\neg\neg(\exists y \in B)(x =_n y) \wedge (\forall y \in B)\neg\neg(\exists x \in A)(x =_n y))$, which is to say $\neg\neg(A =_{n+1} B)$.

■

But actually this is pointless beco's $\neg\neg(A = B)$ implies $\neg\neg(A =_1 B)$ as follows

$\neg\neg(A = B)$

$\neg\neg(\forall x \in A)(x \in B)$

$\neg\neg(\forall x \in A)(\exists y \in B)(x = y)$

$\neg\neg(\forall x \in A)(\exists y \in B)\neg\neg(x = y)$

and the other half too, of course.

We need to drop the outermost $\neg\neg$ from the definition. But would we then be able to prove that $\mathcal{V}$ is $=_n$-unique? I s'pose not.

Is that what we want? Suppose we define the $=_n$ without the $\neg\neg$, so that $x =_{n+1} y$ iff $(\forall x' \in x)(\exists y' \in y)(x' =_n y') \wedge (\forall y' \in y)(\exists x' \in x)(x' =_n y')$. This is a metalanguage recursion over all concrete naturals as before.

We can then prove that $\neg\neg(A =_n B)$ for every n.

Perhaps we can show the following. Let $\sim$ be a homogeneous reflexive relation, and suppose A and B are two Nfinite sets whose double complements are both V. I think we can prove $\neg\neg(A \sim^+ B)$.

$A \sim^+ B$ iff two conjuncts, one of which is

$(\forall x \in A)(\exists y \in B)(x \sim y)$.

Let $x \in A$ be arbitrary. Sse $\neg(\exists y \in B)(x \sim y)$. This implies $\neg(x \in y)$ since $\sim$ is reflexive and x would be a witness. But $\neg\neg(x \in B)$ by assumption on B. So $\neg\neg(\exists y \in B)(x \sim y)$. By UG we now infer $(\forall x \in A)\neg\neg(\exists y \in B)(x \sim y)$, and by Linton-Johnstone we can pull the negneg out to get $\neg\neg(\forall x \in A)(\exists y \in B)(x \sim y)$.

and the other conjunct similarly, giving us $\neg\neg(A \sim^+ B)$.

■

This gives us more than we get from $\neg\neg(A = B)$ beco's that is a congruence relation for stable relations only, whereas $\sim^+$ might not be stable. But we do need it to be stratified, beco's of the induction inside the proof of Linton-Johnstone.

So what is the intersection of all double complements of + of reflexive relations?

The first relation is a congruence relation for all stable relations. What are the later $=_n$ congruence relations for??

Ah, i think that is the right question. $\neg\neg =$ is a congruence relation for all stable relations The point is that R^+ might not be stable even if R is.

However we can prove that if A and B are etc etc then $\neg\neg(\forall x \in A)(\exists! y \in B)(\neg\neg(x = y) \wedge (\forall y \in B)(\exists! x \in A)(\neg\neg(x = y)))$

Next we recall a theorem of Johnstone and Linton from [27] which can be spiced up to prove:

Theorem 42 *If X is subfinite then*

$$(\forall x \in X)\neg\neg\phi \longleftrightarrow \neg\neg(\forall x \in X)\phi$$

holds for stratified ϕ.

Proof:

One direction is easy: constructively $\neg\neg\forall$ implies $\forall\neg\neg$ but not *vice versa*, as we have noted. So let us fix a stratified formula ϕ and prove by induction on X that if X is Kfinite then

$$(\forall y \in X)(\neg\neg\phi(y)) \to \neg\neg(\forall y \in X)(\phi(y)) \qquad (A)$$

(A) is certainly true if $X = \emptyset$. Now assume it true for X, and assume also that $(\forall y \in X \cup \{x\})(\neg\neg\phi(y))$. This last assumption implies both

(i): $(\forall y \in X)(\neg\neg\phi(y))$ and

(ii) $(\forall y \in \{x\})(\neg\neg\phi(y))$,

and (ii) of course implies $\neg\neg\phi(x)$. By induction hypothesis (i) implies

(ii)$'$: $\neg\neg(\forall y \in X)(\phi(y))$.

Now $(\forall y \in X)(\phi(y))$ and $\phi(x)$ together imply

(iii) $(\forall y \in X \cup \{x\})(\phi(y))$

so the conjunction of their double negations will imply the double negation of (iii), namely:

$$\neg\neg(\forall y \in X \cup \{x\})(\phi(y))$$

as desired.

However we claim this also for *subfinite* X. (This fact is not in [27]).

Suppose $X \subseteq A$ where A is Kfinite, and $(\forall x \in X)(\neg\neg\phi(x))$. This is the same as $(\forall x \in A)(x \in X \to \neg\neg\phi(x))$, which implies $(\forall x \in A)\neg\neg(x \in X \to \phi(x))$. By commutation ($A$ is Kfinite) we infer $\neg\neg(\forall x \in A)(x \in X \to \phi(x))$ and thence $\neg\neg(\forall x \in X)\phi(x)$. ■

Where have we used the fact that ϕ is stratified? We need ϕ to be stratified because otherwise the induction we are performing over the Kfinite sets is not stratified.

Can X be a subfinite proper class? No: we need '$x \in X$' to be stratified.

This is certainly true:
$(\forall x \in \mathcal{V})[(\exists y)(y \in x) \to \neg\neg(\exists y \in \mathcal{V})(y \in x)]$
whence
$(\forall x \in \mathcal{V})\neg\neg[(\exists y)(y \in x) \to (\exists y \in \mathcal{V})(y \in x)]$
and, by L-J
$\neg\neg(\forall x \in \mathcal{V})[(\exists y)(y \in x) \to (\exists y \in \mathcal{V})(y \in x)]$
whence it will be consistent that
$(\forall x \in \mathcal{V})[(\exists y)(y \in x) \to (\exists y \in \mathcal{V})(y \in x)]$
but we're not there yet.

REMARK 43
Any two kfinite sets with the same double complement are notnot-equal.

Proof:
If A is kfinite and B is any set then Linton-Johnstone tells us

$$(\forall x \in A)(\neg\neg x \in B) \to \neg\neg(\forall x \in A)(x \in B)$$

So: if A and B are both kfinite we have

$$(\forall x \in A)(\neg\neg x \in B) \to \neg\neg(\forall x \in A)(x \in B)$$

and

$$(\forall x \in B)(\neg\neg x \in A) \to \neg\neg(\forall x \in B)(x \in A)$$

So suppose A and B are both kfinite, with $\sim\sim A = \sim\sim B$. Then $(\forall x \in A)\neg\neg(x \in B)$ whence $\neg\neg(\forall x \in A)(x \in B)$ by Linton-Johnstone. Similarly $(\forall x \in B)(\neg\neg x \in A)$ whence $\neg\neg(\forall x \in B)(x \in A)$. So we have both

$\neg\neg(\forall x \in A)(x \in B)$ and $\neg\neg(\forall x \in B)(x \in A)$. Now $\neg\neg p \wedge \neg\neg q \to \neg\neg(p \wedge q)$ so we have

$\neg\neg[(\forall x \in A)(x \in B) \wedge (\forall x \in B)(x \in A)]$.

Now by extensionality we infer $\neg\neg(A = B)$ ■

I think this means that the intersection al all dense Nfinite sets is a dense Nfinite set. Let X be the intersection of all dense Nfinite sets. Let x be any set, and Y a dense Nfinite set. Then $\neg\neg(x \in Y)$ so, by UG, $(\forall$ dense Nfinite $Y)\neg\neg(x \in Y)$. So $\neg\neg(x \in X)$; so, by UG, $(\forall x)\neg\neg(x \in X)$. So X is dense. And is it Nfinite? It's certainly subfinite.

Isn't every subfinite subset of an Nfinite set Nfinite? I think that is easy to prove by Nfinite induction.

Let's check this properly. Write $D(x)$ and $\mathbb{N}(x)$ for x being dense and Nfinite.

We want $(\forall x)\neg\neg(x \in \bigcap\{y : D(y) \wedge \mathbb{N}(y)\})$

$(\forall x)\neg\neg(\forall y)(D(y) \wedge \mathbb{N}(y). \to x \in y)$

and now the $\neg\neg$ is the wrong side of the $\forall y$.

But it does mean that the class of dense Nfinite sets is closed under Nfinite intersection. Indeed the class of dense sets is closed under Nfinite intersection.

Suppose there is a dense Nfinite set. Remark 43 tells us that any such set is $\neg\neg$ =-unique, so it doesn't much matter which one we consider. So let $\mathcal{V}$ be the intersection of them all. It probably doesn't matter.

For all x and y in $\mathcal{V}$, $\neg\neg(x \cap y \in \mathcal{V})$

(and it's the same for $x \cup y$, $V \setminus x$ and $V \setminus y$) That is to say $(\forall x, y \in \mathcal{V})(\neg\neg(x \cap y \in \mathcal{V}))$

So, by Linton-Johnstone,

$$\neg\neg(\forall x, y \in \mathcal{V})(x \cap y \in \mathcal{V})$$

and of course similarly for all the other algebraic operations. So. for consistency purposes, we may assume that $\mathcal{V}$ really is closed under these operations. Indeed it is going to be not-not a model for SF, stratified foundations.

Better check that it is a notnot model of extensionality.

Suppose x and y are in $\mathcal{V}$, and $(\forall z \in \mathcal{V})(z \in x \longleftrightarrow z \in y)$. We want $x = y$.

Suppose $\neg(w \in x \longleftrightarrow w \in y)$. Such a w is notnot in $\mathcal{V}$. But $\mathcal{V}$ contains no such things. So $\neg(\exists w)\neg(w \in x \longleftrightarrow w \in y)$. But $\neg\exists\neg$ implies $\forall\neg\neg$ so

$(\forall w)\neg\neg(w \in x \longleftrightarrow w \in y)$.

This implies $\neg\neg(x = y)$ So we have proved No!!

$$(\forall x, y \in \mathcal{V})((\forall z \in \mathcal{V})(z \in x \longleftrightarrow z \in y) \to \neg\neg(x = y)).$$

Now we want to appeal to $(A \to \neg\neg B) \to \neg\neg(A \to B)$ (which i'm pretty sure is correct) to get

$$(\forall x, y \in \mathcal{V})(\neg\neg(\forall z \in \mathcal{V})(z \in x \longleftrightarrow z \in y) \to (x = y)).$$

and then we use Linton-Johnstone to get

$$\neg\neg(\forall x, y \in \mathcal{V})((\forall z \in \mathcal{V})(z \in x \longleftrightarrow z \in y) \to (x = y)).$$

[this next paragraph will come to life if we ever show that $\mathcal{V}$ is a model of extensionality]

So $\mathcal{V}$ is notnot a model of extensionality. But it's Nfinite so it obeys excluded middle for atomics, so notnot(it's a model of classical NF). I think that is going to be impossible. I think the fact that $\mathcal{V} \models NF$ forces the logic to be classical. [that will need spelling out!!]

Suppose there is no dense Nfinite set. That means that every Nfinite set has nonempty complement: for every Nfinite x $\neg(\forall y)\neg\neg(y \in x)$. But $\neg\forall\neg\neg$ implies $\neg\neg\exists\neg$ so it also means that for every Nfinite x $\neg\neg(\exists y)(y \notin x)$. Is that enough to show that every Nfinite cardinal notnot has an Nfinite successor? I think so. Now suppose (with a view to obtaining a contradiction) that the set of Nfinite cardinals is Nfinite. Then we can use Linton-Johnstone to show that notnot(every Nfinite cardinal has a successor). But if every Nfinite cardinal has an Nfinite successor then the set of Nfinite cardinals is not Nfinite. So notnot(every Nfinite cardinal has a successor) implies notnot(the set of Nfinite cardinals is not Nfinite). So the set of Nfinite cardinals was not Nfinite after all. Then we have to show that if the collection of Nfinite cardinals is not Nfinite then every Nfinite cardinal has a successor. That sounds obvious but it may be hard.... We can say that the set of cardinals below an nfinite cardinal is nfinite, so if the set of nfinite cardinals it not nfinite it is not the set of cardinals below a given cardinal.

Or perhaps the idea is to show by induction on the Nfinite sets that every one is equipollent to an initial segment of the Nfinite cardinals. Suppose X is Nfinite, and equipollent with some initial segment I of the Nfinite cardinals,

$X \cup \{x\}$ injects into any initial segment I' that properly extends I. so Consider the intersection of all them. Or perhaps we consider $S``I \cup \{0\}$. But then we need to know that S is everywhere defined and we are back where we started.

I think we argue that the set of Nfinite cardinals is discrete (any two Nfinite cardinals are equal or notequal).

Boise march 2019

The challenge seems to be to show that the following are equivalent:

(i) We can interpret Heyting Arithmetic;

(ii) Every $\mathbb{N}$finite cardinal has a successor;

(iii) The set of $\mathbb{N}$finite cardinals is not $\mathbb{N}$finite.

They certainly should be!!

In any case we have $\neg\neg$(there is a dense Nfinite set $\vee$ every Nfinite set has nonempty complement)

The first disjunct should give us Con(NF) or something like it. The second gives us an implementation of Heyting arithmetic.

This looks as if it might come in useful

LEMMA 44
Suppose A and B are two kfinite sets of kfinite sets, and $\sim\sim A \cap kfin = \sim\sim B \cap kfin$.

Then $\sim\sim A = \sim\sim B$.

Proof:
Suppose $x \in \sim\sim A$, which is to say $\neg\neg x \in A$. Then $\neg\neg$ kfin(x), beco's everything in A is kfinite. So $\neg\neg[x \in A \wedge$ kfin(x)]. Then $\neg\neg(x \in \sim\sim B)$. But then $x \in \sim\sim B$, giving $\sim\sim A \subseteq \sim\sim B$. The other inclusion is analogous.

■

Using remark 43 we conclude that if A and B are two kfinite sets of kfinite sets with $\sim\sim A \cap$ kfin $= \sim\sim B \cap$ kfin then $\neg\neg(A = B)$.

So the class of hereditarily kfinite sets (over $\neg\neg =$) is a model of extensionality, as follows. If A and B are hereditarily kfinite sets with the same kfinite $\neg\neg$ members, then (by the above) they have the same double complement, so they are notnotequal.

Other nice things happen. All sets are Nfinite, so equality is decidable. It's a model for power set, because the set of kfinite subsets of a kfinite set is kfinite.

Thinking aloud, as usual (8/vi/2016). We have $\mathcal{V}$ a dense Nfinite set.

First thing to notice is that it is a model for $(\forall xy)(x = y \vee x \neq y)$. This is beco's it is Nfinite, and we prove it by induction. However we can do better.

Consider $\mathcal{V} \cap \{x : (\forall a \in \mathcal{V})(a \in x \vee a \notin x)\}$. It is subfinite and therefore obeys Linton-Johnstone, and its double complement is V, so it is notnot equal to $\mathcal{V}$. Is it Nfinite? I suppose it might not be. Call it $\mathcal{V}'$, and consider $\mathcal{V}' \cap \{x : (\forall a \in \mathcal{V})(a \neq x \vee a \neq x)\}$. This is subfinite and discrete and i think that's enuff to make it Nfinite.

Try to turn $\mathcal{V}$ into a model of classical NF plus not-AxInf. For any ϕ consider $\{x \in \mathcal{V} : \phi^{\mathcal{V}}\}$. The superscript means we have restricted all the parameters and all the bound variables to $\mathcal{V}$. There is something in $\mathcal{V}$ that is notnotequal to this object. That is to say $(\exists a \in \mathcal{V})\neg\neg(\forall x)(x \in a \longleftrightarrow (x \in \mathcal{V} \wedge \phi^{\mathcal{V}}(x)))$.

but we can export '$\neg\neg$' past $\exists$ to get

$\neg\neg(\exists a \in \mathcal{V})(\forall x)(x \in a \longleftrightarrow (x \in \mathcal{V} \wedge \phi^{\mathcal{V}}(x)))$.

and drop the $\neg\neg$ without endangering consistency to obtain

$(\exists a \in \mathcal{V})(\forall x)(x \in a \longleftrightarrow (x \in \mathcal{V} \wedge \phi^{\mathcal{V}}(x)))$.

Doesn't that do it? Well, we have to check extensionality(!)

So sse $x_1 \cap \mathcal{V} = x_2 \cap \mathcal{V}$. Suppose $\neg\neg(z \in x_1)$ Then $(\exists z' \in \mathcal{V})\neg\neg(z = z' \wedge z' \in x_1)$. For this z' we also have $\neg\neg(z' \in x_2 \wedge z' = z)$ so we get $\neg\neg(z \in x_2)$ which is to say that $(\forall z)(\neg\neg(z \in x_1) \longleftrightarrow \neg\neg(z \in x_2))$ which is to say $\sim\sim x_1 = \sim\sim x_2$. That's nice, but what we actually want is $\neg\neg(x_1 = x_2)$ because that would imply $x_1 = x_2$.

Can't we assume x_1 and x_2 are both in $\mathcal{V}$...?

We will need lemma 3 again.

This might help

REMARK 45 *Suppose $Kfin(\{x : \neg\neg\phi(x)\})$. Then it is consistent to assume $(\forall x)(\neg\neg\phi(x) \to \phi(x))$*

Proof:

Write 'A' for '$\{x : \neg\neg\phi(x)\}$) and assume $Kfin(A)$. Then we have $(\forall x \in A)\neg\neg\phi(x)$. Linton-Johnstone now gives us $\neg\neg(\forall x \in A)\phi(x)$ and we can drop the '$\neg\neg$' without imperiling consistency, getting $(\forall x \in A)\phi(x)$ which is to say $(\forall x)(\neg\neg\phi(x) \to \phi(x))$. ■

But are there any such ϕ? $x = \emptyset$ is one. However it is probably the only one: $x = \{\emptyset\}$ is clearly not one!

My guess is that no stable set other than $\emptyset$ is kfinite unless the logic is classical. It might be worth providing a proof. Presumably it proceeds by showing how to map onto Ω any stable set with two distinct elements.

LEMMA 46

Johnstone's weak de Morgan principle—$\neg p \vee \neg\neg p$—implies the analogue of Linton-Johnstone for $\exists$:

$$\neg\neg(\exists x \in A)F(x) \;\to\; (\exists x \in A)(\neg\neg F(x))$$

Proof:

We do this by induction on A. This is all right when $A = \emptyset$. Let's try the induction step. Assume

$$\neg\neg(\exists x \in A \cup \{y\})F(x) \tag{(*)}$$

and aspire to deduce

$$(\exists x \in A \cup \{y\})\neg\neg F(x)$$

The assumption (*) is equivalent to $\neg\neg[(\exists x \in A)(F(x)) \vee F(y)]$.
Using our new-found distributivity (lemma 3)this gives
$\neg\neg[(\exists x \in A)(F(x))] \vee \neg\neg F(y)]$.
Now we use the induction hypothesis on the left disjunct to obtain
$(\exists x \in A)(\neg\neg F(x)) \vee \neg\neg F(y)]$.
which is

$$(\exists x \in A \cup \{y\})(\neg\neg F(x))$$

as desired.

■

Corollary 47
If T is a constructive theory that contradicts classical logic then $T \vdash \neg Kfin(\Omega)$.

Proof:
Let T be a constructive theory that contradicts classical logic. Then

$$T \vdash \neg(\forall x \in \Omega)(x = \bot \vee x = \top) \qquad (B)$$

However, by theorem 38, we have

$$(\forall x \in \Omega)\neg\neg(x = \bot \vee x = \top)).$$

Now, using Johnstone-Linton—and assuming that Ω is finite—we infer

$$\neg\neg(\forall x \in \Omega)(x = \bot \vee x = \top)).$$

which is the negation of (B). So $T \vdash \neg\mathrm{Kfin}(\Omega)$. ■

There now follow a number of observations about Kuratowski-finite sets which are elementary to prove, and well-known to people who are familiar with this material, but probably not to most NFistes. Having struggled to prove them at Peter Johnstone's knee I cannot now resist the temptation to inflict them on the reader.

Remark 48 *Every Kfinite set is empty or inhabited.*

Proof: This is because the collection of sets that are either empty or inhabited contains $\emptyset$ and is closed under adjunction. ■

Remark 49 *Every determinate inhabited subset of x is a quotient of x.*

Proof: If y is a determinate inhabited subset of x then send every member of y to itself and every member of $x \setminus y$ to some arbitrary member of y. ■[6]

[6]It would be nice if this were instead: every determinate nonempty subset .... Do we know that determinate nonempty sets are inhabited?

Remark 50 *Every surjective image of a Kfinite set is Kfinite.*

Proof: We do this by induction. The collection of sets all of whose surjective images are Kfinite contains $\emptyset$ and is closed under insertion. ■

Notice that the surjection in question doesn't have to be a set, as long as it's setlike. In particular:

Corollary 51 *x is kfinite iff $\iota``x$ is finite.*

Remark 52 *If V is Kfinite so is Ω.*

Proof: The function $\lambda x.(x \cap \{\emptyset\})$ maps V onto $\mathcal{P}(\{\emptyset\})$. ■

Lemma 53 *If there is a surjection from A to B there is a surjection from $\mathcal{P}_{kfin}(A)$ to $\mathcal{P}_{kfin}(B)$.*

Proof: Let f be a surjection from A to B. We prove by induction on the (nonempty) Kfinite subsets of B that they are all surjective images of Kfinite subsets of A under f. True for the empty set. Let $B' \cup \{b\}$ be a Kfinite subset of B. By induction hypothesis B' is $f``A'$ for some $A' \subseteq A$ and in any case $b$ is $f(a)$ for some $a \in A$ so $B' \cup \{b\}$ is $f``(A' \cup \{a\})$ as desired. ■

Remark 54 *The set of Kfinite subsets of a Kfinite set is Kfinite.*

Proof:

I am indebted to Peter Johnstone for explaining much of this to me. Let us try to prove this by induction, and see what we need. The empty set has only one subset. How many Kfinite subsets does a singleton have? Two, whatever the size of Ω. Now suppose $\mathcal{P}_{kfin}(x)$ is Kfinite. Let us try to show $\mathcal{P}_{kfin}(x \cup \{y\})$ is finite. We know that $\mathcal{P}_{kfin}(x)$ and $\mathcal{P}_{kfin}(\{y\})$ are finite by induction hypothesis. So $\mathcal{P}_{kfin}(x) \times \mathcal{P}_{kfin}(\{y\})$ is finite too. We know by remark 50 that every quotient of a Kfinite set is Kfinite, so it suffices to show that $\mathcal{P}_{kfin}(x \cup \{y\})$ is a surjective image of $\mathcal{P}_{kfin}(x) \times \mathcal{P}_{kfin}(\{y\})$. This is very far from obvious. Notice that $\mathcal{P}_{kfin}(x) \times \mathcal{P}_{kfin}(\{y\})$ is naturally the same size as $\mathcal{P}_{kfin}(x \sqcup \{y\})$. There is obviously a surjection from $x \sqcup \{y\}$ to $x \cup \{y\}$. Lemma 53 tells us there is a surjection from $\mathcal{P}_{kfin}(x \sqcup \{y\})$ to $\mathcal{P}_{kfin}(x \cup \{y\})$. ■

(Peter Johnstone tells me that this is true because $\mathcal{P}_{kfin}$ is the free semilattice functor.)

Lemma 55
A union of Kfinitely many Kfinite sets is Kfinite;
A product of Kfinitely many Kfinite sets is Kfinite;
If A and B are Kfinite, so is $A \to B$.

Proof: An easy induction.

For the second part we first have to show that the cartesian product of two kfinite sets is kfinite.

It might also be an idea to show that a union of Nfinitely many Nfinite sets is Nfinite, and a product of Nfinitely many Nfinite sets is Nfinite.

For the third part we prove by induction on A that, for all B, if A and B are kfinite, so is $A \to B$. If $A \to B$ is kfinite and B is kfinite, then $\{a\} \to B$ is kfinite whence $(A \to B) \times (\{a\} \to B) = (A \cup \{a\}) \to B$ is also kfinite. Ditto Nfinite of course. ■

Remark 56 *The set of partitions of a kfinite set is not reliably kfinite*

Proof:

Clearly each partition of a Kfinite set is kfinite. By the third part of 55 if there are only kfinitely many partitions then their sumset (which is the power set) will be Kfinite. ■

But perhaps the set of kfinite partitions, or the set of partitions into kfinite pieces, will be kfinite.

Remark 57 *If Ω is Kfinite then the power set of a Kfinite set is Kfinite.*

Proof: Obviously we do this by induction. Base case easy. Now assume $\mathcal{P}(X)$ is Kfinite and deduce that $\mathcal{P}(X \cup \{y\})$ is Kfinite. $\mathcal{P}(X) \times \mathcal{P}(\{y\})$ is Kfinite because $\Omega = \mathcal{P}(\{y\})$ is Kfinite and the product of two Kfinite sets is Kfinite.

$\mathcal{P}(X) \times \mathcal{P}(\{y\})$ is naturally the same size as $\mathcal{P}(X \sqcup \{y\})$. There is a surjection $X \sqcup \{y\} \twoheadrightarrow X \cup \{y\}$ and in general if $f : A \twoheadrightarrow B$ is a surjection, then f lifts to a surjection $\mathcal{P}(A) \twoheadrightarrow \mathcal{P}(B)$. (This is rather in the spirit of lemma 53). So $\mathcal{P}(X \cup \{y\})$ is a quotient of a Kfinite set and is therefore Kfinite. ■

Remark 58 *Every Nfinite family has a choice function*

The classical proof works. Doesn't work for kfinite beco's we might add something that isn't sufficiently distinct from something already in the kfinite family.

Remark 59 *Every Nfinite set is discrete.*

Proof: An easy induction.

Remark 60 *Every Kfinite set is not-not Nfinite.*

One proves this allegation by kfinite induction.

It's true for the empty set, so consider the kfinite set $X \cup \{x\}$ where X is Kfinite. We wish to show that this set is notnot Nfinite. The induction hypothesis is of course that $\neg\neg\text{Nfin}(X)$. With an eye to a *reductio* let us suppose that $X \cup \{x\}$

is *not* Nfinite. So it cannot be the case that $\mathrm{Nfinite}(X) \wedge x \notin X$, beco's that conjunction would imply $\mathrm{Nfinite}(X \cup \{x\})$. So we have $\neg(\mathrm{Nfinite}(X) \wedge x \notin X)$.

If $x \in X$ then $X \cup \{x\} = X$ and $X \cup \{x\}$ is notnot Nfinite as desired;

If $x \notin X$ then $\mathrm{Nfin}(X) \to \mathrm{Nfin}(X \cup \{x\})$, so certainly $\neg\neg(\mathrm{Nfin}(X) \to \mathrm{Nfin}(X \cup \{x\}))$, giving $\neg\neg\mathrm{Nfin}(X) \to \neg\neg\mathrm{Nfin}(X \cup \{x\})$ and the antecedent of this is the induction hypothesis, so we infer the conclusion, namely $\neg\neg\mathrm{Nfin}(X \cup \{x\})$.

So, as long as we have $x \in X \vee x \notin X$, we can conclude $\neg\neg\mathrm{Nfin}(X \cup \{x\})$:

$$(x \in X \vee x \notin X) \to \neg\neg\mathrm{Nfin}(X \cup \{x\}).$$

The conclusion of this conditional is negative, so we can also infer it from the double negation of the antecedent:

$$\neg\neg(x \in X \vee x \notin X) \to \neg\neg\mathrm{Nfin}(X \cup \{x\})$$

Now the antecedent of this last conditional is a constructive thesis, so we infer

$$\neg\neg\mathrm{Nfin}(X \cup \{x\})$$

■

Armed with these concepts we can start thinking about proving the axiom of infinity. One thing we can see almost at once.

Remark 61 *iNF* $\vdash V$ *is not Kfinite.*

Proof:

This is an almost immediate corollary of remark 52, which says that if V is Kfinite then Ω is Kfinite too. But that—as we see in the proof of theorem 38—is enough to imply $\neg\neg(\forall x \in \Omega)(x = \bot \vee x = \top)$. Now $(\forall x \in \Omega)(x = \bot \vee x = \top)$ is enough to prove that V is Dedekind-infinite. So its double negation will prove that V is not-not-Dedekind infinite. And that is enough to show that V cannot be Kfinite: by induction no Kfinite set can be Dedekind-infinite. ■

Recall that a set is subfinite if it has a superset that is Kfinite. since V is not Kfinite, it not subfinite either, and *vacuously* so, since V has no proper supersets at all: its Kfiniteness is a sufficient condition for its subfiniteness. However one cannot run the same argument for $\iota``V$, although that clearly shouldn't be subfinite either. Happily $\iota``V$ is indeed not subfinite, though we do have to do a bit of work to show it.

Remark 62 $\iota``V$ *is not subfinite.*

Proof: Let x and y be two singletons. Then $\neg\neg(x = y \vee x \neq y)$. But x and y were arbitrary, whence

$$(\forall x, y \in \iota``V)\neg\neg(x = y \vee x \neq y).$$

Now suppose that $\iota``V$ were subfinite. Then $\neg\neg$ and $(\forall x \in \iota``V)$ would commute, so we get

$$\neg\neg(\forall x, y \in \iota``V)(x = y \vee x \neq y).$$

and next

$$\neg\neg(\forall x, y)(x = y \vee x \neq y).$$

Now

$$(\forall x, y)(x = y \vee x \neq y) \tag{A}$$

implies that the logic is classical (see remark 64), and thence implies all the theorems of NF, such as: $\iota``V$ is not subfinite. So $(A)$ implies that $\iota``V$ is not subfinite. But then this also follows from $\neg\neg A$. So $\iota``V$ was not subfinite. ■

Naturally the same goes for $\iota^n``V$ for any concrete n.

9.1 Some thoughts about kfiniteness

Reflect that the sumset of a kfinite family of kfinite sets is kfinite. (That was the third part of lemma 55). We might be able to use this to show that certain things are not kfinite.

Let K be a kfinite set, and consider the function $\lambda x.(x \cap K)$. This is a (boolean algebra?) homomorphism from $V = \mathcal{P}(V) \twoheadrightarrow \mathcal{P}(K)$. This surjection partitions V into preimages $\{y : y \cap K = x\}$, one for each $x \subseteq K$. The kernel of this map is $\{x : x \cap K = \emptyset\}$. Does this kernel map onto each $\{y : y \cap K = x \cap K\}$? If so then we can prove that if K is kfinite then $\{x : x \cap K = \emptyset\}$ is not kfinite. If the kernel were kfinite so too would be all the other pieces and then V would be a union of a family of kfinite things and would be kfinite. The trouble is that the family is indexed not by K (which is kfinite) but by $\mathcal{P}(K)$ which might not be. So we can't exploit the fact that a union of kfinitely many kfinite fibres is kfinite.

But let's try anyway. So let's see if, for each x, we can map $\{x : x \cap K = \emptyset\}$ onto $\{y : y \cap K = x \cap K\}$

So fix $x_0 \in \{y : y \cap K = x_0 \cap K\}$. For any other $x \in \{y : y \cap K = x_0 \cap K\}$ we have $x \cap K = x_0 \cap K$. Of course we want $x \,\Delta\, x_0$ to be in the kernel ...but this is easy. Suppose $y \in x \,\Delta\, x_0$. Then clearly $y \notin K$, so $x_0 \,\Delta\, x$ is disjoint from K and is in the kernel. Is this the direction we need...?

Suppose $y \cap K = \emptyset$. We wish to show that $y \,\Delta\, x_0$ belongs to the same preimage as does x_0. So we want

$$(y \,\Delta\, x_0) \cap K = x_0 \cap K. \tag{1}$$

There are various ways of unpacking $y \,\Delta\, x_0$ but the best one is as $(x_0 \cup y) \cap (\overline{x_0} \cup \overline{y})$. So the LHS of equation (1) becomes

$$(x_0 \cup y) \cap K \cap (\overline{x_0} \cup \overline{y}).$$

Now $y \cap K = \emptyset$ so the third intersectand is a superset of K, which means we can ignore it [**is this constructively safe?**] leaving $(x_0 \cup y) \cap K$. But, again, $y \cap K = \emptyset$ so we are left with $x_0 \cap K$ as desired.

What about other homomorphisms? What about $x \mapsto \bigcap\{y : \neg\neg y = x\}$? Does iterating it always reach a fixed point? It's homogeneous! What does the set of fixed points satisfy?

Recycle or delete the rest of this section

If V is Kfinite so is Ω.

Suppose the logic fails to be classical, in the sense that

$$\neg(\forall x)(x \subseteq \{\emptyset\} \to x = \emptyset \vee x = \{\emptyset\}) \tag{BAD}$$

Every Nfinite set has either 0, 1, 2 or more than 2 elements. So if Ω is Nfinite it has precisely two elements.

Ω having precisely two elements implies $(\forall x)(x \subseteq \{\emptyset\} \to x = \emptyset \vee x = \{\emptyset\})$, which contradicts BAD, whence $\neg$ BAD. But every Kfinite set is notnot Nfinite, so we can infer $\neg$ BAD from the weaker assumption that Ω (and a fortiori V) is kfinite. So, if V is kfinite, we can infer

$$\neg\neg(\forall x)(x \subseteq \{\emptyset\} \to x = \emptyset \vee x = \{\emptyset\}).$$

But that means that we can consistently add $(\forall x)(x \subseteq \{\emptyset\} \to x = \emptyset \vee x = \{\emptyset\})$ to the theory we are working in, which is *i*NF + Kfin(V).

But if we do add this then the logic becomes classical and we can run Specker's proof of $\neg$AxInf. So if *i*NF + kfin(V) were consistent so too would *i*NF + classical logic + Kfin(V) be consistent. But it isn't. So *i*NF proves V is not kfinite.

There doesn't seem to be any doubt that one can prove by Nfinite induction that

REMARK 63 *Every Nfinite set is either*
(i) empty or
(ii) has precisely one member or
(iii) has precisely two members or
(iv) has at least three distinct members.

Proof:
The empty set satisfies (i). Thereafter consider $X \cup \{x\}$ with X Nfinite, and $x \notin X$.

If X satisfies (i) then $X \cup \{x\}$ satisfies (ii):
if X satisfies (ii) then $X \cup \{x\}$ satisfies (iii):
if X satisfies (iii) or (iv) then $X \cup \{x\}$ satisfies (iv).

X must satisfy one of them, by induction hypothesis, so $X \cup \{x\}$ must satisfy one as well. So that seems OK. ■

There doesn't seem to be any doubt that if Ω is Nfinite it must have precisely two members. It is a constructive thesis that it cannot have three distinct members, and it clearly has at least two. And if it has precisely two then the logic is classical.

Let us suppose the logic is not classical, so we have something like

$$\neg(\forall x \subseteq \{\emptyset\})(x = \emptyset \vee x = \{\emptyset\}).$$

Then Ω does not have precisely two members, so it is not Nfinite. But if it's not Nfinite it can't be kfinite either, beco's kfinite implies notnot Nfinite.

9.2 Duality and Double Duality

Duality is the scheme $\phi \longleftrightarrow \hat{\phi}$ for all ϕ, where $\hat{\phi}$ is the result of replacing $\in$ by $\notin$ throughout ϕ. Classically duality is provable for weakly stratified ϕ (although its status for unstratified formulæ is obscure); intuitionistically it is strong.

Classically the dual of an axiom of NF is an axiom of NF. Well, logically equivalent to an axiom of NF modulo some pretty basic set theory. So the dual of a theorem of NF is a theorem of NF. This is not true constructively: the dual of extensionality is $(\forall xy)(\sim\sim x = \sim\sim y \to x = y)$ which is very strong, and presumably not a theorem of *i*NF (tho' there is no proof of this at present). So don't expect *i*NF $\vdash \phi$ iff *i*NF $\vdash \hat{\phi}$. And we should not be surprised if it turns out that adding the duality scheme $\phi \longleftrightarrow \hat{\phi}$ to *i*NF gives classical NF or something like it. And it does so turn out.

Remark 64 *The following are equivalent*
1. $\forall xy(x = y \vee x \neq y)$*;*
2. $\forall xy(x \in y \vee x \notin y)$*;*
3. All singletons have precisely two subsets;
4. (Universal closure of) excluded middle for weakly stratified formulæ;
5. All subsets of singletons are Kfinite;
6. $(\forall xy)(\neg\neg x = y \to x = y)$*;*
7. $(\forall xy)(\neg\neg x \in y \to x \in y)$*;*
8. Duality for stratified formulæ;
9. Double duality for all formulæ.
10. Ω *is subfinite.*

Proof:

We will prove: $1 \to 2$; $2 \to 1$; $1 \to 3 \wedge 5$; $4 \to 1 \wedge 2$; $5 \to 3$; $3 \to 5$; $7 \to 2$; $6 \to 1$; $8 \vee 9 \to 6 \wedge 7$; $7 \to 6$. Must incorporate 10!

$1 \to 2$.

Assume 1. This tells us that

$\{z : z = a \wedge x \in y\} = \{a\}$ or $\{z : z = a \wedge x \in y\} \neq \{a\}$.

The first possibility implies $x \in y$ and the second $x \notin y$.

$2 \to 1$

Either $y \in \{x\}$ or $\neg(y \in \{x\})$. In the first case $y = x$ by $\in$-elimination. In the second $y \neq x$ by $\in$-elimination.

$1 \to 3 \wedge 5$.

Assume 1 and let x be a subset of a singleton $\{a\}$. Then $x = \{a\} \vee x \neq \{a\}$. In the first horn x is finite. In the second horn, it must be the case that not everything in $\{a\}$ is in x. But then if a were in x we would have $x = \{a\}$. So $a \notin x$. So x is empty. Either way x is finite. And there are only these two possibilities. So we infer 3 and 5.

$4 \to 1 \wedge 2$

Both 1 and 2 are special cases of 4. That is to say: *i*NF + excluded middle for weakly stratified formulæ is the same as *i*NF + excluded middle for atomic formulæ. Since intuitionistic **Z** + excluded middle for atomics is the same as **Z** one might think that *i*NF + excluded middle for weakly stratified formulæ is simply NF but this is not so. This is because we do not have comprehension for unstratified formulæ in *i*NF. In **Z** the proof of $3 \to 4$ that we have here proves excluded middle for *all* formulæ.

$3 \to 4$.

Let ϕ be an arbitrary weakly stratified formula whose free variables are to be found in $\vec{z}$, and 'x' a variable not free in it. Then $\{x : x = V \wedge \phi\}$ is a set by weakly stratified comprehension and

$$(\forall \vec{z})(\{x : x = V \wedge \phi\} = \{V\} \ \vee \ \{x : x = V \wedge \phi\} = \emptyset),$$

since $\{V\}$ has only two subsets, itself and $\emptyset$.

$5 \to 3$

Every Kfinite set is either empty or inhabited, by remark 48. Let x be a subset of a singleton a. By 5, x is finite, so it is either inhabited, in which case it is equal to a, or it's the empty set. So a has only two subsets.

$3 \to 5$.

If $\{a\}$ has precisely two subsets, they must be $\emptyset$ and $\{a\}$, both of which are finite.

$7 \to 2.$

(Daniel Dzierzgowski showed me how to do this). For any x and y we have $\neg\neg(y \in x \vee \neg(y \in x))$. That is to say, $\neg\neg(y \in \{z : z \in x \vee \neg(z \in x)\})$. By 7 this implies $y \in \{z : z \in x \vee \neg(z \in x)\}$ and thence $y \in x \vee \neg(y \in x)$. The other direction is easy.

$6 \to 1$

Suppose $\{x : x = a \wedge p\} \neq \{a\}$ and $\neg\neg p$. The first assumption gives us $\neg p$, which contradicts $\neg\neg p$. This proves $\neg\neg\{x : x = a \wedge p\} = \{a\}$, and thence (by (6)) $\{x : x = a \wedge p\} = \{a\}$ which implies p. So (6) implies $\neg\neg p \to p$.

$(8 \vee 9) \to (6 \wedge 7).$

The way to derive 6 and 7 from 8 and 9 is to notice that the dual and double dual of extensionality are $(\forall xy)(x = y \longleftrightarrow (\forall z)(z \notin x \longleftrightarrow z \notin y))$ and $(\forall xy)(x = y \longleftrightarrow (\forall z)(\neg\neg z \in x \longleftrightarrow \neg\neg z \in y))$. Each of these implies that every set is equal to its closure (that is to say $x = {\sim}{\sim}x$) and is therefore stable.

For the converse notice that if every set is stable then double duality holds without restriction; if complementation is 1-1 and onto then the usual argument proves duality for weakly stratified formulæ.

$7 \to 6.$

Assume $(\forall xy)(\neg\neg x \in y \to x \in y)$ and $\neg\neg(u = v)$. By extensionality we have $\neg\neg(\forall z)(z \in u \longleftrightarrow z \in v)$. Now suppose $x \in u$. If $x \notin v$ we would derive a contradiction, so $\neg\neg(x \in v)$ whence $x \in v$. But x was arbitrary, so $(\forall x)(x \in u \to x \in v)$. Similarly $(\forall x)(x \in v \to x \in u)$, and $u = v$ as desired.

$6 \to$ excluded middle for weakly stratified formulæ.

$$
\begin{prooftree}
\AxiomC{$[\{x : \phi \vee \neg\phi\} \neq V]^1$}
\RightLabel{$\in$-elim}
\UnaryInfC{$\neg(\forall x)(\phi \vee \neg\phi)$}
\AxiomC{$[\phi \vee \neg\phi]^2$}
\RightLabel{$\forall$-int}
\UnaryInfC{$(\forall x)(\phi \vee \neg\phi)$}
\RightLabel{$\to$-elim}
\BinaryInfC{$\perp$}
\RightLabel{$\to$-int (2)}
\UnaryInfC{$\neg(\phi \vee \neg\phi)$}
\AxiomC{$\neg\neg(\phi \vee \neg\phi)$}
\RightLabel{$\to$-elim}
\BinaryInfC{$\perp$}
\RightLabel{$\to$-int (1)}
\UnaryInfC{$\neg\neg(\{x : \phi \vee \neg\phi\} = V)$}
\RightLabel{Double Negation}
\UnaryInfC{$\{x : \phi \vee \neg\phi\} = V$}
\RightLabel{$\in$-elim}
\UnaryInfC{$\phi \vee \neg\phi$}
\end{prooftree}
\tag{2}
$$

$6 \to$ double negation for weakly stratified formulæ

Double negation for atomics implies double negation for weakly stratified formulæ

$$
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[x=a\wedge p]^2}{x=a}\,\wedge\text{-elim}}{((x=a)\wedge p)\to x=a}\,\to\text{-int (1)} \qquad \dfrac{\dfrac{[p]^3 \qquad [x=a]^2}{x=a\wedge p}\,\wedge\text{-int}}{x=a\to x=a\wedge p}\,\to\text{-int (2)}}{x=a\longleftrightarrow x=a\wedge p}\,\wedge\text{-int}}{(\forall x)(x=a\longleftrightarrow x=a\wedge p)}\,\forall\text{-int}}{\{x:x=a\}=\{x:x=a\wedge p\}}\,\in\text{-int} \qquad [\{x:x=a\}=\{x:x=a\wedge p\}\to\bot]^4}{\bot}\,\to\text{-elim}}{p\to\bot}\,\to\text{-int (3)} \qquad \neg\neg p^5}{\bot}
$$

$$
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\bot}{\neg\neg(\{x:x=a\}=\{x:x=a\wedge p\})}\,\to\text{-int (4)}}{\{x:x=a\}=\{x:x=a\wedge p\}}\,\neg\neg\text{-cancel}}{(\forall x)(((x=a)\wedge p)\longleftrightarrow(x=a))}\,\in\text{-elim}}{((a=a)\wedge p)\longleftrightarrow(a=a)}\,\forall\text{-elim}}{(a=a)\to p}\,\wedge\text{-elim} \qquad \cdots
$$

$$
\dfrac{p}{\neg\neg p\to p}\,\to\text{-i}\cdots \tag{3}
$$

The compiler seems not to like this proof

If we know that complementation is 1-1 we can apply the preservation theorem for permutations to infer 8, duality for weakly stratified formulæ. Complementation being 1-1 follows from $(\forall x)(x = \ {\sim}{\sim}x)$.

We can infer 9 if we know $x\in y \longleftrightarrow \neg\neg(x\in y)$, because then we can use substitutivity of the biconditional.

■

Lemma 65 *Double negation for atomics implies double negation for all formulæ built up from atomics by $\wedge$, $\neg$, $\vee$ and $\forall$.*

Proof: We prove the lemma by structural induction. (Notice that since this proof does not use comprehension it will hold for *all* formulæ in the range of the negative interpretation not just all *stratified* formulæ in the range of the negative interpretation.)

$\bigwedge$ For the induction assume $A\vee\neg A$ and $B\vee\neg B$. Then, by distributivity, we have

$$(A\wedge B)\vee(A\wedge\neg B)\vee(\neg A\wedge B)\vee(\neg A\wedge\neg B).$$

The last three disjuncts all imply $\neg(A\wedge B)$ so we infer

$$(A\wedge B)\vee\neg(A\wedge B)$$

as desired.

$\bigvee$ For the induction assume $A\vee\neg A$ and $B\vee\neg B$. Then, by distributivity, we have

$$(A\wedge B)\vee(A\wedge\neg B)\vee(\neg A\wedge B)\vee(\neg A\wedge\neg B).$$

The first three disjuncts all imply $A \vee B$ and the last disjunct implies $\neg(A \vee B)$ so we infer

$$(A \vee B) \vee \neg(A \vee B)$$

as desired.

$\neg$ is easy.

$\forall$ We assume $F(a) \vee \neg F(a)$ and deduce $(\forall x)(F(x)) \vee \neg(\forall x)(F(x))$.

$$
\begin{prooftree}
\AxiomC{$[\forall x F(x)]^1$}
\RightLabel{$\forall$-elim}
\UnaryInfC{$F(a)$}
\AxiomC{$[\neg F(a)]^2$}
\RightLabel{$\to$-elim}
\BinaryInfC{$\bot$}
\RightLabel{$\to$-int (1)}
\UnaryInfC{$\neg(\forall x)(F(x))$}
\AxiomC{$[\neg\neg(\forall x F(x))]^3$}
\RightLabel{$\to$-elim}
\BinaryInfC{$\bot$}
\RightLabel{$\to$-int (2)}
\UnaryInfC{$\neg\neg F(a)$}
\AxiomC{$(\forall x)(\neg\neg(F(x)) \to F(x))$}
\RightLabel{$\forall$-elim}
\UnaryInfC{$(\neg\neg F(a)) \to F(a)$}
\RightLabel{$\to$-elim}
\BinaryInfC{$F(a)$}
\RightLabel{$\forall$-int}
\UnaryInfC{$\forall x F(x)$}
\RightLabel{$\to$-int (3)}
\UnaryInfC{$\neg\neg(\forall x F(x)) \to \forall x F(x)$}
\end{prooftree}
\quad (4)
$$

$\to$

$$
\begin{prooftree}
\AxiomC{$[A]^3$}
\AxiomC{$[A \to B]^1$}
\RightLabel{$\to$-elim}
\BinaryInfC{B}
\AxiomC{$[\neg B]^2$}
\RightLabel{$\to$-elim}
\BinaryInfC{$\bot$}
\RightLabel{$\to$-int (1)}
\UnaryInfC{$\neg(A \to B)$}
\AxiomC{$\neg\neg(A \to B)$}
\RightLabel{$\to$-elim}
\BinaryInfC{$\bot$}
\RightLabel{$\to$-int (2)}
\UnaryInfC{$\neg\neg B$}
\AxiomC{$\neg\neg B \to B$}
\RightLabel{$\to$-elim}
\BinaryInfC{B}
\RightLabel{$\to$-int (3)}
\UnaryInfC{$A \to B$}
\end{prooftree}
\quad (5)
$$

■

In particular it holds for all formulæ in the range of the negative interpretation. Notice that the induction doesn't work for $\exists$. We can obtain a forcing/Kripke countermodel for $\forall x(F(x) \vee \neg(F(x)) \vdash \exists x F(x) \vee \neg\exists x F(x))$ in which there are two worlds. Make the root world empty and put some frogs in the second world.

Notice that we cannot prove double duality for stratified formulæ—not even for *closed* stratified formulæ—unless we make at least *some* extra assumptions: think about $(\forall x)(x = {\sim}{\sim}x)$. This is stratified and quite strong, but its double dual is a tautology!

This probably shows that double duality for stratified formulæ is as strong as double duality for all formulaæ.

And notice that this is true *despite* excluded middle for closed formulæ being consistent wrt *i*NF!

Does Duality lead to a Consistency Proof?

Logical duality is the operation of swapping atomic formulae with their negations. Classical propositional logic is self-dual: the dual of a tautology is a tautology. Constructive logic not so. There is a duality scheme for set theory that is the scheme of biconditionals $\phi \longleftrightarrow \widehat{\phi}$ where $\widehat{\phi}$ is the dual of ϕ (though we do not negate equations). In NF the instances of this scheme that are stratified are theorems; it is conjectured that the full scheme is consistent relative to NF but this has not been shown. In the constructive setting this duality scheme is of course strong and (suitably phrased) gives us classical logic. I wrote this up in [21] so we don't really need it here.

Suppose you have a possible world model of some constructive set theory, that is to say a structure of signature $\in, =$. Define a new model by keeping the old worlds, the old accessibility relation and the old equality, but now say that a world W in the new structure believes $x \in y$ iff W believed $\neg(x \in y)$ under the old dispensation. (We will have to check that this new model obeys persistence ...!) Notice that this means that the new structure will satisfy excluded middle (for atomics) if the old model satisfies $\neg p \vee \neg\neg p$ for atomics.

What can one say about this new structure? Suppose the old structure believed ϕ, a formula satisfying the rather odd property that every occurrence of '$\in$' has a slash through it. (I hope it will become clear why this is less crazy than it sounds). Then the new model satisfies the modification of ϕ obtained by removing all those slashes.

Let's check persistence. Suppose $W \models x \in y$. That means that in the old model no W' beyond W believed $x \in y$. So, for every W'' beyond W and every W' beyond W'', $W' \not\models x \in y$ ... and, in fact, $W' \models x \notin y$—which is to say that $W' \models x \in y$ in the new model. I am assuming that the induction steps for the connectives behave themselves, but that should be checked.

What I am after is a possible world structure with the feature that when you wave this particular wand over it you get a model of *i*NF. However the analysis that I am going to wade through is not really very sensitive to a choice of comprehension scheme. All i need is that we have comprehension for some set Γ of formulæ s.t. whenever ϕ is in Γ then the *dual* of ϕ (put a slash through every occurrence of '$\in$') is also in Γ.

What must such a structure look like, and can we find one?

Of course it only has to satisfy a *restricted* (or perhaps *modified*) version of comprehension: $(\forall\vec{x})(\exists y)(\forall z)(z \notin y \longleftrightarrow \phi(\vec{x}, z))$ where 'ϕ' is weakly stratifiable (or rather in Γ, *mutatis mutandis*) and all occurrences of '$\in$' have slashes through them—and that *sounds* like something one might be able to do something with. However you want the new structure to satisfy extensionality and that means that the original structure has to satisfy

$$(\forall xy)(x = y \longleftrightarrow (\forall z)(z \notin x \longleftrightarrow z \notin y)) \qquad \text{(Beefed-up Extensionality)}$$

On the face of it this looks a lot stronger than ordinary extensionality since it says that two sets with the same double complement are equal, and that looks as if it will enforce the law of double negation and make our logic classical. But, as it turns out, life is a bit more complicated than that.

If we are to think of this system proof theoretically it has an inference rule for beefed-up extensionality, and an introduction and elimination rule for $\notin$—not for $\in$. This is going to have the effect that altho' there are going to be plenty of proofs of things like $t \notin t'$, it is going to be difficult if not impossible for the last line of any proof to be $t \in t'$. The hope is that this just might save our bacon, and give rise to a proof-theoretic demonstration of consistency: the theory might be consistent for silly proof theoretical reasons.

Let's see what this theory (the one that when dualised gives *i*NF) looks like.

Observe that in an instance of the comprehension scheme

$$(\forall\vec{x})(\exists y)(\forall z)(z \notin y \longleftrightarrow \phi(\vec{x}, z))$$

any witness to the '$\exists y$' must be unique—by beefed-up extensionality—so there is no escape into NFU down that route.

Does every set have a complement?

$$(\forall x)(\exists y)(\forall z)(z \notin y \longleftrightarrow \neg(z \notin x))$$

is something like an axiom of complementation, and it is an axiom of comprehension of the appropriately restricted kind. We can reason about the (unique!) y satisfying $(\forall z)(z \notin y \longleftrightarrow \neg(z \notin x))$ as follows. Let's give it the suggestive name 'x^*'.

Let z be arbitrary. Suppose $z \in x^*$; then, by contraposing the R $\to$ L implication, we infer $\neg\neg(z \notin x)$ whence $z \notin x$. But what about the other direction? Easy to show that some things are notnot in x^*, but there doesn't seem to be any way of concluding that anything actually *is* in x^*! But at least x and x^* are disjoint.

However we can do better than that: * is involutive and injective.

Lemma 66
$(\forall a)(a^{**} = a)$*;*
$(\forall a, b)(a^* = b^* \to a = b)$.

Proof: For the first part it will suffice to show $(\forall x)(\neg(x \in a^{**}) \longleftrightarrow \neg(x \in a))$. Fix a. Let x be arbitrary; we have $\neg(x \in a^*) \longleftrightarrow \neg\neg(x \in a)$ by comprehension. This implies the result of negating both sides:

$\neg\neg(x \in a^*) \longleftrightarrow \neg\neg\neg(x \in a)$ which of course is $\neg\neg(x \in a^*) \longleftrightarrow \neg(x \in a)$.

Similarly comprehension gives us

$\neg(x \in a^{**}) \longleftrightarrow \neg\neg(x \in a^*)$. This time we do not need to negate both sides; all we need to do is compose the biconditionals ... eliminating $\neg\neg(x \in a^*)$ to get $\neg(x \in a^{**}) \longleftrightarrow \neg(x \in a)$ as desired. But x was arbitrary, whence $(\forall x)(\neg(x \in a^{**}) \longleftrightarrow \neg(x \in a))$ as desired. Now we use beefed-up extensionality to infer $a^{**} = a$.

The second part (injectivity) follows from involutiveness. $a^* = b^*$ gives $a^{**} = b^{**}$. But then $a = a^{**} = b^{**} = b$.

■

Remark 67 *Stability of equality*
$(\forall a, b)(\neg\neg(a = b) \to a = b)$.

Proof:
Let a and b be arbitrary, with $\neg\neg(a = b)$. Using substitutivity of equality, $\neg\neg(a = b)$ and $F(a)$ implies $\neg\neg F(b)$, so take F to be $z \notin a$, z arbitrary. This gives $z \notin a \to \neg\neg z \notin b$ and the consequent is of course $z \notin b$. We get the other direction analogously, whence $(\forall z)(z \notin a \longleftrightarrow z \notin b)$, and we can use beefed-up extensionality on this to infer $a = b$.

■

There doesn't seem to be an analogous argument to show that $\in$ is stable (there being no obvious way to exploit BUE).

Let's think about $B(x)$ in this context. We have an axiom

$(\forall x)(\exists y)(\forall z)(z \notin y \longleftrightarrow x \notin z)$ and, for each x, this y is unique by BUE. Call this object $B(x)$ (a bit of overloading)

We want to show that the collection of the Bs is a model of the classical theory. Suppose $(\forall x)(B(x) \notin B(a) \longleftrightarrow B(x) \notin B(b))$. We want $B(a) = B(b)$.

$B(x) \notin B(a)$ iff $a \notin B(x)$ iff $x \notin a$, and b similarly, so $B(a)$ and $B(b)$ have the same nonmembers so are identical by BUE.

Thus the collection of Bs equipped with the negation of $\in$ satisfy extensionality on the nose. What sort of comprehension does it satisfy?

If i give you $B(x)$ can you recover x? At least in the sense that $B(x) = B(y) \to x = y$?

$B(x) = B(y) \to$

???

10 Interpreting Arithmetic

[thinking aloud]

We want

$$(\forall x)(Nfin(x) \to (\exists y)(y \notin x))$$

but we'd settle for

$$\neg\neg(\forall x)(Nfin(x) \to (\exists y)(y \notin x)).$$

So let's attempt to prove it by *reductio*. Assume

$$\neg(\forall x)(Nfin(x) \to (\exists y)(y \notin x)).$$

But we can't do anything with $\neg\forall$; we can do something with $\neg\forall\neg\neg$ co's that implies $\neg\neg\exists\neg$

So let's try to prove

$$\neg\neg(\forall x)\neg\neg(Nfin(x) \to (\exists y)(y \notin x))$$

instead. So assume

$$\neg(\forall x)\neg\neg(Nfin(x) \to (\exists y)(y \notin x))$$

which will imply

$$\neg\neg(\exists x)\neg(Nfin(x) \to (\exists y)(y \notin x))$$

which implies

$$\neg\neg(\exists x)(Nfin(x) \wedge \neg(\exists y)(y \notin x))$$

So we would be able to consistently add

$$(\exists x)(Nfin(x) \wedge \neg(\exists y)(y \notin x))$$

This would be an Nfinite set whose double complement is V. There doesn't seem to be anything preventing that.

But might such a set give us a classical model of NF?

We have seen (remark 61) that V cannot be Kfinite. In the classical case this is enough to provide us with an implementation of arithmetic: if there is a set X that is not actually inductively finite, then it has subsets of all inductively finite sizes and $\mathcal{P}^2(X)$ will contain cardinals of all those sets (in the form of their local equipollence classes), and therefore a copy of $\mathbb{N}$. In the constructive setting life is a great deal more complicated. For one thing, a set can fail to be inductively finite for silly reasons. Every Kfinite set is either empty or inhabited, so any nonempty uninhabited set fails to be finite. There can even be subsets of singletons with this property. Clearly sets like this are not going to give rise to implementations of Heyting Arithmetic. Another concern is that the set that is to be the set of Heyting naturals has to be discrete. This means that if we are to try to implement the naturals as the cardinals of Xfinite sets for some idea Xfinite of finiteness, then we must be able to prove that any two Xfinite sets are either the same size or not the same size. The situation is complex, but the best candidate for Xfiniteness is Nfiniteness.

It is of course true that we are not, in principle, constrained to implement Heyting naturals as equipollence classes of Xfinite sets. All one needs is a countable discrete set equipped with suitable operations. However it is easy to show that if there is such a set then its initial segments furnish us with Nfinite sets of all the requisite sizes and we could have implemented our Heyting naturals as equipollence classes after all. This is cleared up in remark 68.

We know that V is not Nfinite, so certainly we can prove

$$(\forall x)(\text{Nfinite}(x) \to \neg(\forall y)(y \in x)) \tag{6}$$

However, if we are to use Nfinite cardinals as our implementation of Heyting arithmetic we would need to know that every Nfinite cardinal has an Nfinite successor, and for that we would need the (potentially much stronger)

$$(\forall x)(\text{Nfinite}(x) \rightarrow (\exists y)(y \notin x)). \tag{7}$$

Although this is of course classically equivalent to (6) the two are not constructively equivalent. It may be worth mentioning the intermediate formula

$$(\forall x)(\text{Nfinite}(x) \rightarrow \neg\neg(\exists y)(y \notin x)). \tag{8}$$

in this connection. It is equivalent to the assertion that there is no dense Nfinite set.

As for implementing Heyting Arithmetic, the following might help to clear the air.

Remark 68 *The following are equivalent:*

1. *Heyting Arithmetic can be implemented in iNF;*
2. *iNF proves the existence of a Dedekind-infinite discrete set;*
3. *The cardinals of Nfinite sets give an implementation of Heyting arithmetic.*

Proof:
$3 \rightarrow 1$ is obvious. $1 \rightarrow 2$ is easy. If there is an implementation of Heyting Arithmetic the set of naturals is Dedekind infinite and discrete.

$2 \rightarrow 1$

> Let X be a set with a 1-1 map $f : X \rightarrow X$ such that $X \setminus (f``X)$ is inhabited, by x_0 say. Consider the inductively defined set
>
> $$\mathbb{N} = \bigcap\{Y : x_0 \in Y \wedge f``Y \subseteq Y\}$$
>
> We prove by induction on x that $(\forall y \in \mathbb{N})(y = x \vee y \neq x)$. True for x_0 because everything in $\mathbb{N}$ is either x_0 or a value of f in which case it is not $= x_0$. Now suppose true for x, we want to infer it for $f(x)$. Think of an arbitrary $y \in \mathbb{N}$. Either $y = x_0$ (in which case the second disjunct is satisfied) or $y = f(z)$ for some $z \in \mathbb{N}$. But this reduces to $x = z \vee x \neq z$ which is true by induction hypothesis.

$1 \rightarrow 3$.

> We prove by induction that every Nfinite set is the size of an initial segment of the naturals. If $\pi : X \rightarrow Y$ is a bijection between X and an initial segment Y of the naturals, then the function
>
> $\lambda x.(\texttt{if}\ \ x \in X\ \ \texttt{then}\ \ S(\pi x)\ \ \texttt{else}\ \ 0)$

maps $X \cup \{y\}$ one-to-one onto an initial segment of $\mathbb{N}$. So the collection of cardinals of Nfinite sets is unbounded.

■

I turned up this fact in the course of my search for results that would give us an implementation of Heyting arithmetic:

Remark 69 *The set of cardinals of Kfinite sets is not subfinite.*

Proof: Suppose it were. Then, by the same sort of use of Johnstone-Linton, we conclude that not-not any two Kfinite sets are the same size or different sizes. This means that we can consistently add the assertion that any two Kfinite sets are the same size or different sizes.

Let x and y be any two sets. $|\{x\}| = |\{x, y\}| \vee |\{x\}| \neq |\{x, y\}|$. $|\{x\}| = |\{x, y\}|$ implies $x = y$ and $|\{x\}| \neq |\{x, y\}|$ implies $x \neq y$. Decidability of equality implies that the logic is classical—for weakly stratified formulæ at least—and that is enough to prove the axiom of infinity. And if the axiom of infinity holds, the set of Kfinite cardinals is definitely not subfinite. ■

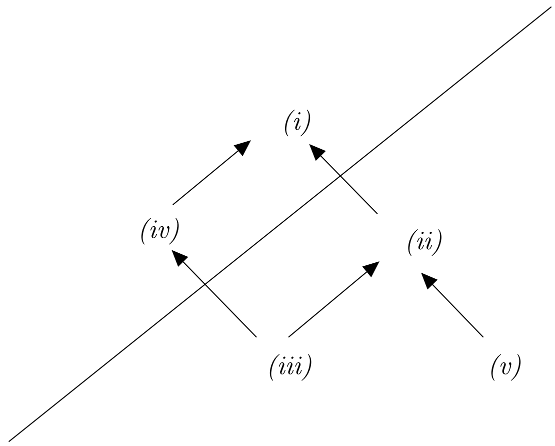
However it is not much use. More useful would be a discovery that the set of Nfinite cardinals is not subfinite, but that doesn't seem to be on offer!

This illustrates a phenomenon in *i*NF that one eventually gets used to if one persists long enough. There are these collections of objects which one would like to be infinite. For example, one would like there to be infinitely many natural numbers, naturally(!) In the *i*NF context, collections like this can turn out to be infinite not for the sound reason that there genuinely are as many of the objects as we desire, but because excluded middle has failed in a big way and the objects one seeks (natural numbers etc) although not numerous, have *multifurcated* and turned into *slop*. Remark 69 is a case in point. It tells us that the collection of cardinals of kfinite sets is not small, but the reason why it is not small is that lots of cardinals which we ought to be able to prove identical we can't. The pile of unexcluded possibilities of equations swells the size of the set of finite cardinals.

Remark 70 *The five propositions:*
(i) $\neg\neg =$ *is an equivalence relation of finite index;*
(ii) $V/\neg\neg =$ *has a transversal;*
(iii) There is a dense discrete set;
(iv) There is a dense Nfinite set;
(v) $\neg A \vee \neg\neg A$ *for* A *stratified.*[7]

are related as in the following pseudo–Hasse-diagram. The arrows indicate increasing strength!!

[7] This is the stratified version of Peter Johnstone's de Morgan principle from [26].



Proof:

(i) implies (ii).

The quotient is discrete, so if it is Kfinite it is Nfinite. The usual classical proof that every inductively finite set of disjoint nonempty sets has a transversal set can easily be modified to give a constructive proof that every Nfinite set of disjoint inhabited sets has a transversal set, which will be a witness to (ii). (Observe that we can also prove constructively by induction that every Nfinite family of inhabited sets has a selection function. Nfinite but not Kfinite!)

(i) implies (iv).

The transversal obtained in the proof of (i) $\to$ (ii) is a witness to (iv).

(iv) implies (iii).

Obvious.

(ii) implies (v).

If there is a transversal then there is a total surjection from V onto it which preserves $\neg\neg =$. We ask what happens to a member of Ω. It must get sent to $\perp$ or to $\top$.

(ii) implies (iii).

The transversal is both dense and discrete.

(iv) $\wedge$ (v) $\to$ (i).

This is the hard part! The key fact here is that (v) implies that $\neg\neg$ commutes with $\exists$ on Kfinite domains (see lemma 46). Let $\mathcal{V}$ be a witness to (iv), a dense Nfinite set. For any x we have $\neg\neg(x \in \mathcal{V})$. Now $(x \in \mathcal{V})$ implies $(\exists z \in \mathcal{V})(x = z)$ so $\neg\neg(x \in \mathcal{V})$ implies $\neg\neg(\exists z \in \mathcal{V})(x = z)$. But by commutativity we then infer $(\exists z \in \mathcal{V})\neg\neg(x = z)$. But x was arbitrary. So everything is not-not-equal to something in $\mathcal{V}$. So $\mathcal{V}$ is a finite transversal set for the quotient $V/\neg\neg =$. But then there is a bijection between $\iota``\mathcal{V}$ and $V/(\neg\neg =)$ (send each singleton to the unique equivalence class in which it is included). Any bijective copy of a kfinite set is kfinite and, for any $x$, $x$ is kfinite iff $\iota``x$ is kfinite by corollary 51 This proves (i).

■

I now think i have a proof that (v) implies (ii).

Remark 71 *Weak de Morgan implies that if A is Nfinite and $x \in \sim\sim A$ then $(\exists x' \in A)\neg\neg(x = x')$.*

Proof:
Base case. Suppose our Nfinite set is $\{a, b\}$. Then $x \in \sim\sim\{a, b\}$ says $\neg\neg(x = a \vee x = b)$, which implies $\neg((x \neq a) \wedge (x \neq b))$. Now we use weak de Morgan (which tells us that $\neg\neg$ distributes over $\vee$) to infer $\neg\neg(x = a) \vee \neg\neg(x = b)$, whence $(\exists x' \in \{a, b\})(\neg\neg(x = x'))$.

For the induction suppose that $x \in \sim\sim A \to (\exists b \in A)(\neg\neg(x = b))$. Now suppose $a \notin A$ and $x \in \sim\sim(A \cup \{a\})$. That is to say $\neg\neg(x \in A \vee x = a)$ which implies $\neg(x \notin A \ \wedge\ x \neq a)$. Now invoke $x \neq a \ \vee\ \neg\neg(x = a)$. The first horn gives $\neg\neg(x \in A)$ (at which point we appeal to the induction hypothesis) and the second horn gives $\neg\neg(x = a)$.

■

Is this enough to prove that V is not the double complement of an Nfinite set? If it were, there would be a finite collection s.t. everything is notnotequal to something in it. Does this give a model of the classical theory?

Let us also make a note of the fact that:

Remark 72
Johnstone's weak de Morgan principle(v) for stratified formulæ is equivalent to the assertion that $V/(\neg\neg =)$ is discrete.

Proof: Assume Johnstone's weak de Morgan principle and let p and q be two elements of $V/(\neg\neg =)$. Then either $p \neq q$ or $\neg\neg(p = q)$ by weak de Morgan. If the second then $p = q$ by the special properties of $V/(\neg\neg =)$: if two $\neg\neg =$-equivalence classes are not-not-equal then they are equal.

For the other direction, assume discreteness of $V/(\neg\neg =)$. Then $[x] \neq [y] \vee \neg\neg([x] = [y])$ for any x and y. So let ϕ be any stratified formula and set $x =: [[\phi]]$ and $y =: \top$. This gives us $[[\phi]] \neq \top \ \vee \ \neg\neg([[\phi]] = \top)$, which is to say $\neg\phi \vee \neg\neg\phi$. ■

Things below and to the right of the diagonal line in the picture on p. 67 we would be happy to believe. Things to the left and above we would not. (i) would imply that NF is not consistent and (iv) implies that there is no implementation of Heyting arithmetic in *i*NF.

If (i) holds, so that $\neg\neg =$ is an equivalence relation of finite index then there will be a transversal set $\mathcal{V}$ for it. $\mathcal{V}$ will be Nfinite, being a discrete surjective image of an Nfinite set. But $\mathcal{V}$ is highly pathological. We know

$$(\forall x)(\neg\neg(x \in \mathcal{V})) \tag{A}$$

but we also know, since $\mathcal{V}$ is Nfinite, that it cannot be equal to V, so—by extensionality—

$$\neg(\forall x)(x \in \mathcal{V}) \tag{B}$$

But the conjunction of (A) and (B) contradicts classical logic.

If (iv) holds, so there is a dense Nfinite set, the cardinal of this set is an Nfinite cardinal lacking a successor. This means that the Nfinite cardinals do not afford us an implementation of Heyting arithmetic. But this means, by remark 68, that there is no implementation of Heyting Arithmetic at all.

H I A T U S

If no Nfinite set has uninhabited complement then we get an implementation of Heyting arithmetic. If we have a Nfinite set whose double complement is V then we *might* get a model of classical NF. What one is really looking for is a dilemma one horn of which gives an implementation of Heyting arithmtic and the other of which gives a model of the classical theory

H I A T U S

Now return to our project of finding an implementation of $\mathbb{N}$ in *i*NF. We have to show that there are no Nfinite sets whose double complement is V.

Consider $V/(\neg\neg =)$. It is a partition. If it is Kfinite then it has a selection set, and that selection set will be a kfinite set whose double complement is V and we don't want *that*!

Suppose $\sim\sim X = V$. Send $x \in X \mapsto [x]_{\neg\neg=}$. Does this map X onto $V/(\neg\neg =)$? It would be good if it did, but i suspect it doesn't. This is because we would seem to need $(\forall y)(\exists x \in X)(\neg\neg(y = x))$ but that doesn't obviously follow from $\sim\sim X = V$. Mind you, its double negation would do:

$$\neg\neg(\forall y)(\exists x \in X)(\neg\neg(y = x)) \ldots$$

Fix X with $\sim\sim X = V$. Send $x \in X \mapsto [x]_{\neg\neg=}$. We want to show that this is notnot onto. If it isn't onto then there is y s.t. $(\forall x \in X)(y \neq x)$. So in particular $y \notin x$. But this contradicts $\sim\sim X = V$. So this map is, indeed, notnot onto. Can we show that there is a map $X \to \Omega$ that is notnot onto? $x \in X \mapsto x \cap \{\emptyset\}$? Suppose there is $p \subseteq \{\emptyset\}$ that is not in the range of this map then certainly $p \notin X$.

Let's go over this slowly....

Suppose $\sim\sim X = V$. Consider $\pi = \lambda x \in X.(x \cap \{\emptyset\})$. We claim that $(\forall p \in \Omega)\neg\neg(\exists x \in X)(\pi(x) = p)$

Let $p \in \Omega$ be arbitrary. Then we have $\neg\neg(p \in X)$. So certainly $\neg\neg(\pi(p) \in \pi``X)$ so $\neg\neg$(p is a value of π). That is to say

$$(\forall p \in \Omega)\neg\neg(\exists x \in X)(p = \pi(x))$$

which is all very well, but the '$\neg\neg$' is in the wrong place; we want:

$\neg\neg(\forall p \in \Omega)(\exists x \in X)(p = \pi(x))$

where the stuff inside the $\neg\neg$ implies that X is not kfinite, so we would infer that X is not kfinite.

If A is a set of stable sets then $(\forall x, y \in A)(\neg\neg(x = y) \to x = y)$. If $x = y$ then $(\forall z)(z \in x \longleftrightarrow z \in y)$. So $\neg\neg(x = y)$ implies $\neg\neg(\forall z)(z \in x \longleftrightarrow z \in y)$. We can push the '$\neg\neg$' inside to get $(\forall z)\neg\neg(z \in x \longleftrightarrow z \in y)$. But this implies $\sim\sim x = \sim\sim y$. But x and y are stable.

A random thought... Suppose $x \in a$. Consider $x' = \{z : z \in x \wedge p\}$. Then $p \to x' \in y$ but perhaps not conversely.

11 Extensions of *i*NF

It is not hard to check that the result of adding to *i*NF all the formulæ of remark 64 is a system as strong as NF. (This was known to Dzierzgowski.)

Remark 73

1. *Commutation of* $\neg\neg$ *and* $\forall$ *implies that* $\sim\sim x = \sim\sim y \to \neg\neg(x = y)$ *and*
2. *this is enough to interpret NF.*

Proof:

1. $\sim\sim x = \sim\sim y \to \neg\neg(x = y)$

   Extensionality tells us that $\sim\sim x = \sim\sim y$ is

$$(\forall z)((\neg\neg(z \in x) \to \neg\neg(z \in y) \wedge (\neg\neg(z \in y) \to \neg\neg(z \in x))))$$

Now constructively $\neg\neg A \to \neg\neg B$ implies $\neg\neg(A \to B)$ and $\neg\neg A \wedge \neg\neg B$ implies $\neg\neg(A \wedge B)$ so we infer

$$(\forall z)\neg\neg(z \in x \longleftrightarrow z \in y)$$

and we can now pull the $\neg\neg$ out by (i) to obtain

$$\neg\neg(\forall z)(z \in x \longleftrightarrow z \in y)$$

which, by extensionality, is $\neg\neg(x = y)$.

2. The interpretation takes the universe to be V, and takes $=$ to be not-not-equality and $\in$ to be not-not-membership. Since $\sim\sim x = \sim\sim y \to \neg\neg(x = y)$ then the equivalence relation of not-not-equality is a congruence relation for not-not-membership and the quotient is extensional and obeys classical logic. ■

However there is a very idiomatic interpretation of NF into *i*NF + commutation-of-$\forall$-with-$\neg\neg$, which works as follows.

Define $\overline{B}x =: \{y : x \notin y\}$. The carrier set of our model $\mathfrak{M}$ will be $\overline{B}$"V, equality will be equality and membership of the model will be $\in$.

The model satisfies comprehension: $\{x : \phi(x, \overrightarrow{\overline{B}(y)})\}$ in the sense of $\mathfrak{M}$ will be $\overline{B}(\{x : \phi(x, \vec{y})\})$. The model satisfies double negation for atomics:

(i) for $\in$ by the following reasoning:

$\overline{B}(x) \in \overline{B}(y)$ is the same as $\neg(y \in \overline{B}(x))$ (by definition of $\overline{B}$). $y \in \overline{B}(x)$ is the same as $x \notin y$ by definition of $\overline{B}$ so $\neg(y \in \overline{B}(x))$ is the same as $\neg\neg(x \in y)$

(ii) Double negation for $=$ follows:

$$\neg\neg(\overline{B}(x) = \overline{B}(y))$$

iff

$$\neg\neg\forall z(z \in \overline{B}(x) \longleftrightarrow z \in \overline{B}(y))$$

iff

$$\neg\neg\forall z(x \notin z \longleftrightarrow y \notin z)$$

Now we we can import the $\neg\neg$ past the $\forall$ and the formula within the scope of the $\forall$ is stable so we get

$$\forall z(x \notin z \longleftrightarrow y \notin z)$$

which is equivalent (by dfn of $\overline{B}$)

$$\forall z(z \in \overline{B}(x) \longleftrightarrow z \in \overline{B}(y))$$

whence $\overline{B}(x) = \overline{B}(y)$ by extensionality. ■

Given that we have double negation for atomics we can now prove that double negation holds for all formulæ in the range of the negative interpretation.

This will give us an interpretation of classical NF *via* the standard negative interpretation as long as the model is extensional.

Observe that

Lemma 74 $\overline{B}(x) = \overline{B}(y)$ *iff* $\neg\neg(x = y)$

Proof:
$\overline{B}(x) = \overline{B}(y)$ implies $(\forall z)(z \in \overline{B}(x) \longleftrightarrow z \in \overline{B}(y))$
whence in particular
$\{x\} \in \overline{B}(x) \longleftrightarrow \{x\} \in \overline{B}(y)$
but the LHS of the biconditional is false, whence
$\{x\} \notin \overline{B}(y)$ whence
$\neg\neg(y \in \{x\})$ whence finally
$\neg\neg(y = x)$

For the other direction we have $\neg\neg(x = y) \to \neg\neg(\overline{B}(x) = \overline{B}(y))$, but we have just seen that $\neg\neg(\overline{B}(x) = \overline{B}(y)) \to \overline{B}(x) = \overline{B}(y)$ ■

$\mathfrak{M}$ believes extensionality as long as

$$(\forall z)(\overline{B}(z) \in \overline{B}(x) \longleftrightarrow \overline{B}(z) \in \overline{B}(y))$$

implies $\overline{B}(x) = \overline{B}(y)$. Now $\overline{B}(z) \in \overline{B}(x)$ is equivalent to $\neg\neg(z \in x)$ so the displayed formula is equivalent to

$$(\forall z)(\neg\neg(z \in x) \longleftrightarrow \neg\neg(z \in y)).$$

This is equivalent to

$$(\forall z)\neg\neg(z \in x \longleftrightarrow z \in y).$$

This is the point at which we use commutation-of-$\forall$-with-$\neg\neg$; we infer

$$\neg\neg(\forall z)(z \in x \longleftrightarrow z \in y).$$

By extensionality this is $\neg\neg(x = y)$, which implies $\neg\neg(\overline{B}(x) = \overline{B}(y))$; and we established in (ii) that equality is stable.

This proves that $\overline{B}$“V is not kfinite. If it were, $\overline{B}$“V would be a model of (classical) NF that believes V to be kfinite.

Notice that for (i)–(iii) we need the full strength of *tertium non datur* for atomics from remark 64; *tertium non datur* for closed formulæ is much weaker.

REMARK 75 *If iNF is consistent, so is iNF + the scheme $\phi \vee \neg\phi$ for all closed ϕ.*

Proof:

If one cannot prove $\neg p$ then one can add p as an axiom. If $\neg p$ is not provable then no contradiction can be deduced from p!

Consider now a conjunction

$$C: \bigwedge_{i<n}(p_i \vee \neg p_i)$$

of expressions of the kind we are considering. By an old result of Glivenko ([23] and [24]) $\neg\neg C$ is a theorem of intuitionistic propositional logic so $\neg C$ cannot be a theorem of intuitionistic propositional logic. By the preceding remark, it must be possible to adjoin C consistently.

By compactness the scheme is now consistent. ■

It now (may 2011) seems to me that one should be able to do a bit more with this. Suppose that $\overline{B}$"V is subfinite, and exploit Johnstone-Linton.

Pick up a random finite tuple $\vec{x}$ of things from $\overline{B}$"V. Let C be a boolean combination of atomic assertions about the various $\vec{x}$, which happens to be a truth-table tautology. Then, by Glivenko's theorem, we must have $(\forall\vec{x} \in \overline{B}``V)\neg\neg C$. But then, by Johnstone-Linton (since $\overline{B}$"$V$ is subfinite), we must have $\neg\neg(\forall\vec{x} \in \overline{B}``V)C$. So we can consistently add $(\forall\vec{x} \in \overline{B}``V)C$ for all such C.

In particular we must be able to add simple things like

$$(\forall xy \in \overline{B}``V)(x = y \vee \neg(x = y))$$

and this of course is equivalent to

$$(\forall xy)(\neg\neg(x = y) \vee \neg(x = y))$$

...the tho'rt being that if we get enuff things like this we could infer the consistency of the classical theory. That way we would prove that if $kfin(\overline{B}``V)$ is consistent with *i*NF then NF is consistent.

Remark 64 has quite a lot to say about excluded middle and suchlike for weakly stratified formulæ. What about unstratified formulæ? The situation was investigated by Daniel Dzierzgowski in the works cited. He noticed that if we can find two structures $\mathfrak{M}$ and $\mathcal{N}$ which are both models of an NF-like theory T such that $\mathfrak{M}$ is a substructure of $\mathcal{N}$ elementary for stratified formulæ but $\mathfrak{M}$ and $\mathcal{N}$ are not elementarily equivalent then we can incorporate these two structures into a Kripke model for an intuitionistic version of T in which excluded middle fails for unstratified formulæ. He challenged the NF istes to find such $\mathfrak{M}$, $\mathcal{N}$ and T. Three examples came up, provided by Friederike Körner and me and so we now know

Remark 76

1. *Intuitionistic NF 0 + term rule for weakly stratified formulæ*[8] *does not prove excluded middle for unstratified formulæ.*

2. *If iNF is consistent it doesn't prove excluded middle for unstratified formulæ.*

3. *If iNF + term rule for weakly stratified formulæ is consistent it doesn't prove excluded middle for unstratified formulæ.*

[8] NF O is the theory whose axioms are extensionality and existence of $\{x : \phi(x, \vec{y})\}$ where ϕ is stratified and quantifier-free.

12 Stable Sets and Negative Interpretations

12.1 Stuff to fit in

Claim: Every bijective copy of a stable set is stable.

Suppose $f : A \to B$ is a bijection, and B is stable: $\sim\sim B = B$.

Write '$A(f, a, b)$' for '$f(a = b)$'.

We have $a \in A \to (\exists b)(A(f, a, b) \wedge b \in B)$ and

$a \in A \to (\forall b)(A(f, a, b) \to b \in B)$

Work inside ./ Suppose $\neg\neg(a \in A)$. We get $\neg\neg(\exists b)(A(f, a, b) \wedge b \in B)$ and $\neg\neg(\forall b)(A(f, a, b) \to b \in B)$ which latter gives us $(\forall b)\neg\neg(A(f, a, b) \to \neg\neg(b \in B))$ but B is stable so we get $(\forall b)\neg\neg(A(f, a, b) \to (b \in B))$

Every subfinite set is notnot Nfinite.

Suppose every subset of X is notnot Nfinite. We want every subset of $X \cup \{x\}$ to be notnot Nfinite.

Suppose $Y' \subseteq X \cup \{x\}$. Then $Y' = Y \cup z$ where $Y = y' \cap X$ and $z = Y \cap \{x\}$. Y is notnot Nfinite by assumption. We want z to be notnot Nfinite. One must avoid the trap of saying that $z \subseteq \{x\}$ and $\{x\}$ is Nfinite so z is notnot Nfinite, co's that is like the proof that all billiard balls are the same colour. We first have to prove that every subset of a singleton is notnot Nfinite. We certainly have $\neg\neg(z = \{x\} \vee z = \emptyset)$. Both $\{x\}$ and $\emptyset$ are Nfinite so z must be notnot Nfinite as desired.

So Y' is a union of two disjoint sets both of which are notnot Nfinite. But a union of two disjoint Nfinite sets is Nfinite. So, reasoning inside the notnots we conclude that Y is notnot Nfinite.

This powers an inductive proof that every subset of every Nfinite set is notnot Nfinite.

We can define $=$ by $x = y$ iff they belong to the same things. If x and y belong to the same stable sets then they are notnot equal. Ther may be something cute and general one can say about equivalence relations x resembles y iff they belong to the same ϕ things. Notice that if x and y belong to the same Nfinite sets then they are identical. This difference is something to do with Nfinite being a set-theoretical property and stability being a logical property. Is the equivlence relation $\sim\sim x = \sim\sim y$ (which appears below) expressible in this style? Yes, trivially. Consider the set X of all equivalence classes under this equivalence relation. Two sets have the same double complement iff they belong to the same members of X. Duh!

Some of these equivalence relations are 1-symmetric. Which tells us that yer typical model of *i*NF has very few permutations. This is for the following reason. Suppose $\neg\neg(x = y)$, and that there is a permutation of V that—whatever else it does—it sends x to u and y to v. Then $\neg\neg(u = v)$.

In fact, suppose $x \sim y$—where $\sim$ is one of these 1-symmetric equivalence relations. Suppose further that there is a permutation that sends x to u and y to v, then $u \sim v$. So, consider the the action of $\mathrm{Symm}(V)$ on pairs; it must have lots of orbits, not just one. To be clear about it, p is a pair iff

$(\exists a, b)(\forall x)(x \in p. \longleftrightarrow .x = a \vee x = b)$. Classically Symm($V$) has two orbits on pairs: the singletons and the pairs. Constructively God knows.

Notice that Linton-Johnstone means that the set of stable sets is a topology!

Truth-values are Sets

The scheme
$(\exists x)(\forall y)(y \in x \longleftrightarrow A)$
for A a closed formula (a sentence)
is consistent relative to *i*NF. No stratification requirement!

Proof:

Fix A and consider the formula $(\forall y)(y \in x \longleftrightarrow A)$. Abbreviate it to $\alpha(x)$. Observe that $\neg\alpha(V) \to \neg A$ and $\neg\alpha(\emptyset) \to \neg\neg A$. It is simple enuff to find a proof of the sequent

$$\neg(\neg F(a) \wedge \neg F(b)) \vdash \neg\neg\exists x F(x)$$

so we conclude $\neg\neg\exists x\alpha(x)$. So we can consistently add $\exists x\alpha(x)$. Can we do this for all A simultaneously? I think so. Suppose

$$\neg(\exists x\alpha(x) \wedge \exists x\beta(x)).$$

This implies

$$\neg\neg(\neg\exists x\alpha(x) \vee \neg\exists x\beta(x)).$$

Now $(\neg\exists x\alpha(x) \vee \neg\exists x\beta(x))$ must lead to a contradiction, since we can prove both $\neg\neg\exists x\alpha(x)$ and $\neg\neg\exists x\beta(x))$. So the double negation leads to a contradiction too. So we deduce

$$\neg(\exists x\alpha(x) \wedge \exists x\beta(x)).$$

And of course we could have done this for any finite number of As.

What happens if A has free variables in it (other than 'x')? We conclude as before that $\neg\neg\exists x\alpha(x)$ except that this time it's $(\forall\vec{u})\neg\neg\exists x\alpha(x)$ where the $\vec{u}$ mops up all the variables free in A. And now the $\neg\neg$ is the wrong side of the quantifier.

12.1.1 The intersection of all dense sets

Call it X. It's 1-symmetric. Is it empty? Inhabited? dense?

Observe that if $x \in X$ then
(A) X and $(X \setminus \{x\}) \cup \{x\}$ have the same double complement
and
(B) $y \in (X \setminus \{x\}) \cup \{x\}) \cap [[x]]_{\neq} \to y = x$.
For (A) suppose $\neg\neg(y \in X)$ but $y \notin ((X \setminus \{x\}) \cup \{x\})$. Then both
(i) $y \notin (X \setminus \{x\})$ and
(ii) $y \neq x$.
(i) gives $\neg(y \in X \wedge \neg\neg(y = x))$. The second conjunct inside the $\neg$ is refuted by (ii) so we infer $\neg(y \in X)$ but this contradicts the assumption that $\neg\neg(y \in X)$.

For (B) Assume the antecedent. Then $y \in (X \setminus \{x\}) \cup \{x\})$. Then either $y = x$ (which is what we want) or $y \in (X \setminus \{x\})$. This implies $y \in X \wedge y \neq x$. But if $y \neq x$ we cannot have $y \in [[x]]_{\neg\neg=}$.

12.2 Double complements and fishy sets might yet prove the axiom of infinity

The idea is to show that if A is an Nfinite set then everything in $\sim\sim A$ is a fishy combination of things in A. This will bound the size of $\sim\sim A$ and make it im possible for there to be a dense Nfinite set. Naturally we will prove this by Nfinite induction.

Start by thinking about $\sim\sim\{a, b\}$, where $a \neq b$.

Suppose $t \in \sim\sim\{a, b\}$. Then t cannot be distinct from both a and b. I claim t is the fishy combination:

$$\{x : (x \in a \wedge t = a) \vee (x \in b \wedge t = b)\}$$

Call this chap `fishy` for short. We want `fishy` to be coextensive with t.

Sse $x \in$ `fishy`. Then $(x \in a \wedge t = a) \vee (x \in b \wedge t = b)$. But both disjuncts separately imply $x \in t$. So `fishy` $\subseteq t$.

The other direction might not work. Let's try. Humph. We seem to get only that $t \subset \sim\sim$`fishy`. We certainly have $x \in t \wedge (t = a \vee t = b). \to x \in$ `fishy`.

That suggests that we haven't quite got the definition of `fishy` right.

No, that's not the moral. The moral is there might be lots of things in $\sim\sim\{a, b\}$. $\sim\sim$distributes over $\cap$ but not over $\cup$. ('co's $\neg\neg$ distributes over $\wedge$ but not $\vee$!)

12.3 Prologue on Negative Interpretations and a Cautionary Tale

The obvious question to think about is whether or not there is a negative interpretation of NF into *i*NF. 'Negative interpretation' covers a multitude of sins, but here is a very discouraging reflection ....

Let us say that a recursively defined map * is a [insert adjective here] interpretation if

(i) ϕ^* is quantifier-free when ϕ is atomic,

(ii) ϕ^* is a constructive thesis whenever ϕ is a classical thesis, and

(iii) If ϕ is closed then $\phi \longleftrightarrow \phi^*$ is a classical thesis.

The closedness condition in (iii) is is to allow $(x \in y)^*$ to be $x \notin y$.

Observe that Holmes' nice interpretation of NFU into *i*NF is not [adjective] because the interpretation of $x \in y$ is not quantifier-free. However any interpretation traditionally regarded as negative will tick these three boxes.

Observe also that the composition of two [adjective] interpretations is another [adjective] interpretation.

So the result is as follows.

Remark 77 *Let T_1 be a classical theory of sets with full extensionality (no atoms) and T_2 a constructive theory of sets. If * is an [adjective] interpretation of T_1 into T_2 satisfying $(x \in y)^* = \neg\neg(x \in y)$ then T_2 proves commutation of $\forall$ and $\neg\neg$ for all formulæ for which it has comprehension.*

Proof:

Any [adjective] interpretation of extensionality more-or-less has to be $(\forall xy)(\sim\sim x = \sim\sim y \to \neg\neg(x = y))$. But this implies commutation of $\forall$ with $\neg\neg$, at least when the stuff inside the quantifier is stuff for which we have comprehension. Work in *i*NF for the sake of an illustration. Suppose $(\forall x)\neg\neg\phi$. This says $\sim\sim\{x : \phi\} = V$. (Here we use comprehension on ϕ.) Applying $(\forall xy)(\sim\sim x = \sim\sim y \to \neg\neg(x = y))$ we get $\neg\neg(\{x : \phi\} = V)$, and *that* is $\neg\neg(\forall x)\phi$. ■

However this result relies on the range of variables in [adjective] interpretations being the whole universe. And that is perhaps excessive.

But there is already a proof of this fact in these notes. There is also a proof that (if we are working in *i*NF) that is enough to interpret NF, by means of $\{\overline{B}y : y \in V\}$. See sections 11 and 14.

It may be worth noting that if T is a constructive theory extending the constructive version of NFO then $T+$ commutatation interprets the classical version of T.

That was done in *i*NF; if we want to do the same when T_2 is something like CZF then we have to do it "locally".

If you add commutation-of-$\forall$-with-$\neg\neg$ to *i*NFU (the constructive version of *i*NF) then you can prove that all empty sets are notnotequal.

I think that commutation implies that notnotequality is a congruence relation for $\neg\neg\in$ and obeys extensionality.

Suppose everything in A is notnotequal to something in B and *vice versa*; commutation will imply $\neg\neg(A = B)$, as follows.

Sse $(\forall x \in A)(\exists y \in B)(\neg\neg(x = y))$. This is
$(\forall x)(x \in A \to (\exists y \in B)(\neg\neg(x = y)))$ whence
$(\forall x)(\neg\neg(x \in A) \to \neg\neg(\exists y \in B)(\neg\neg(x = y)))$ and
$(\forall x)(\neg\neg(x \in A) \to \neg\neg(\exists y \in B)(x = y))$ and
$(\forall x)(\neg\neg(x \in A) \to \neg\neg(x \in B))$ and the other direction too of course.

Commutation now gives $\neg\neg(A = B)$. Commutation also implies the negative interpretation of extensionality.

So i think this means that *i*NFU + commutation interprets NF.

Does Randall's interpretation of NFU in *i*NF give us an interpretation of NFU into *i*NFU?

[when did i ask this question?? Anyway, the answer appears to be YES]

12.4 Stable Sets

Stability is an interesting notion because of the obvious possible connections with negative interpretations. The facts which follow might be useful to the reader who is trying to get a feel for what is going on: as we remarked above, $\sim\sim$is obviously order-preserving, inflationary, and idempotent. It behaves a bit like a closure operator. Stable sets are sets fixed by $\sim\sim$. The first thought is that the collection of hereditarily stable sets should always be a model for the classical version of whatever set theory is in hand. It turns out that this is not the case, but watching how it goes wrong can be quite enlightening for the beginner. So: let's collect some facts about stable and hereditarily stable sets. As far as possible this discussion will be done in a minimal set theory so we don't have to worry whether it's NF or ZF.

The following easy fact is beginning to look hugely important:

Lemma 78 *If $\neg\neg(x = y)$ and $\sim\sim x = x$ and $\sim\sim y = y$ then $x = y$.*

Proof:

$z \in x$ and $\neg\neg(x = y)$ imply $\neg\neg(z \in y)$, but $y = \sim\sim y$ so $z \in y$. The opposite inclusion is analogous. ■

This gives an injection from the set of stable sets (or rather the set of singletons of stable sets) into the quotient $V/(\neg\neg =)$. And $V/(\neg\neg =)$ is a subset of the set of stable sets—every equivalence class is stable. So these cardinals are tied closely together:

$$T|\{x : \text{stab}(x)\}| \;\leq\; |(V/(\neg\neg =))| \;\leq\; |\{x : \text{stab}(x)\}|$$

In corollary 18 we prove that $V/(\neg\neg =)$ is not finite. This is good, beco's the bottom level of $\mathfrak{M}$ is $V/(\neg\neg =)$, and we want $\mathfrak{M}$ to be a model of AxInf.

Lemma 79 *The set of stable sets is closed under arbitrary intersection.*

Proof:

Let A be a set of stable sets; we wish to prove $\sim\sim\bigcap A = \bigcap A$. To that end suppose $x \in \sim\sim\bigcap A$. That is to say $\neg\neg(x \in \bigcap A)$ which is $\neg\neg(\forall a)(a \in A \rightarrow x \in a)$. We can push $\neg\neg$ inside thru' $\forall$ to obtain $(\forall a)\neg\neg(a \in A \rightarrow x \in a)$. We can push $\neg\neg$ again to obtain $(\forall a)(\neg\neg(a \in A) \rightarrow \neg\neg(x \in a))$ whence certainly $(\forall a)((a \in A) \rightarrow \neg\neg(x \in a))$. But now every $a \in A$ is stable so we get $(\forall a)(a \in A \rightarrow x \in a)$, which is to say $x \in \bigcap A$. That was on the assumption of $x \in \sim\sim\bigcup A$, so we have proved $(\forall x)(x \in \sim\sim\bigcup A \rightarrow x \in \bigcap A)$ which is of course $\sim\sim\bigcap A \subseteq \bigcap A$. The inclusion in the other direction is obvious, so we conclude $\sim\sim\bigcap A = \bigcap A$.

■

A lot follows from closure under arbitrary intersection. Because triple negation is the same as single negation, $V \setminus x$ is always stable, whence it follows that $x \setminus y$ is always stable as long as x is: we don't even need y to be stable!

12.4.1 Flat Sets

It also means that the collection of stable sets is a complete poset under inclusion, this despite the fact that a union of two stable sets might not be stable. The meet of two stable sets x and y is not $x \cup y$ but $\sim (\sim x \wedge \sim y)$ ... it's got to contain fishy combinations of things in x with things in y.

Not only is it a complete poset, it is complemented: $V \setminus x$ is stable if x is, and is the complement of x in the sense of the poset. I thought briefly that the set of stable sets might be finite but there are actually lots of stable sets. Whenever x is a stable set and 'y' is not free in p then $\{y \in x : \neg\neg p\}$ is stable. In particular, if 'y' is not free in p then $\{y : \neg\neg p\}$ is stable.

DEFINITION 80 *Let us say a set S is flat iff $(\forall xy)(x \in S \longleftrightarrow y \in S)$.*

Sets of the form $\{y : p\}$ where 'y' is not free in p are flat.

And the converse is true too. A flat set x is equal to $\{y : x = V\}$.

The function $x \mapsto x \cap \{\emptyset\}$ maps the set of flat sets onto Ω. Let $p \in \Omega$. Consider the flat set $\{x : \emptyset \in p\}$. Then $\{x : \emptyset \in p\} \cap \{\emptyset\} = \{x : x = \emptyset \wedge \emptyset \in p\} = p$. This gives us a bijection between Ω and the flat sets.

LEMMA 81 *The power set of a stable set is stable.*

Proof:

Suppose x is stable. We want $X = \{y \subseteq x : stab(y)\}$ to be stable. So we want

$$y \in \ \sim\sim X \to y \in X.$$

Assuming $y \in \ \sim\sim X$ we have

$$\neg\neg(y \subseteq x)$$

This is

$$\neg\neg(\forall z)(z \in y \to z \in x).$$

$\neg\neg$ can be pushed inside $\forall$ so we get

$$(\forall z)\neg\neg(z \in y \to z \in x)$$

which certainly implies

$$(\forall z)(z \in y \to \neg\neg(z \in x)).$$

But x is stable so we get

$$(\forall z)(z \in y \to z \in x)$$

as desired. ∎

Notice that this does not prove that the set of stable subsets of a stable set is stable. I don't think the set of stable sets is stable. Notnotstable $\not\to$ stable. It would be nice to show that this implies a strong logical principle.

This is probably the correct place to collect some facts about double complement. $\sim\sim(A \cap B) = \ \sim\sim A \cap \ \sim\sim B$. This is true beco's $\neg\neg$ distributes over $\wedge$. This means, for example, that $\{y \subseteq A : \sim\sim y = \ \sim\sim A\}$ is a filter on A!

12.4.2 Hereditarily stable sets

With an eye to negative interpretations for set theory one naturally asks: "What *hereditarily* stable sets are there?" Obviously $\emptyset$ at least! Looking further out ...

- $\{\emptyset\}$ is stable—and therefore hereditarily stable—beco's $\neg\neg(y \in \{\emptyset\})$ is $\neg\neg(y = \emptyset)$ which is $\neg\neg(\forall x)(x \notin y)$ which implies $(\forall x)(x \notin y)$ which is to say $\sim\sim\{\emptyset\} = \{\emptyset\}$.

- Can a Quine atom be stable? If it is, then of course it is hereditarily stable. If a is a stable Quine atom then $(\forall b)(\neg\neg(a = b) \rightarrow a = b)$. That looks possible, but our interest in hereditarily stable sets is—in the first instance—primarily in the wellfounded version.

- $\{\{\emptyset\}\}$ looks as if it ought to be hereditarily stable doesn't it? But consider: if it is, we can do the following. Let y be a "dense" truth-value (truth values are subsets of $\{\emptyset\}$, the generic singleton, and a truth value is dense if its double negation (its double complement) is the `true` (is $\{\emptyset\}$)) so that $\neg\neg(y = \{\emptyset\})$. Then $y \in \sim\sim\{\{\emptyset\}\}$ so $y \in \{\{\emptyset\}\}$ since $\{\{\emptyset\}\}$ is stable by assumption. So $y = \{\emptyset\}$. So all dense truth-values are actually equal to the `true`. This is the principle of double negation isn't it?

So the thought is that if there are any [wellfounded] stable sets of stable sets other than $\emptyset$ and $\{\emptyset\}$ then logical principles follow, specifically double negation.

OK, so $\{\{\emptyset\}\}$ being a stable set of stable sets has consequences. How about its double complement; can that be a stable set of stable sets? Suppose it were. Then

$y \in \sim\sim\{\{\emptyset\}\} \rightarrow y$ stable.

The antecedent is equivalent to $\neg\neg(y = \{\emptyset\})$. So every dense truth value is a stable set. But every such truth value is a set notnotinhabited by the empty set. So the empty set genuinely inhabits it. So it is equal to the true.

Again we get a logical principle.

We will need the following

Lemma 82 *Every stable set contains all fishy combinations of its members.*

Proof: Let X be a stable set, with a, b both in X. Consider the fishy combination $f = \{x : (x \in A \wedge p) \vee (x \in B \wedge \neg p)\}$. If *per impossibile* $f \notin X$ then $f \neq a$ and $f \neq b$, but f cannot be distinct from both a and b so we infer $\neg\neg(f \in X)$. But X was stable, whence $f \in X$. ■

The upshot seems to be that the correct inductive structure for use in relative consistency proofs in the negative interpretation style is not the hereditarily stable sets but might be the structure formed by adding at each stage double complements of all subsets. This leads us to a definition due originally to Powell, [28] in 1975.

12.5 Models of TST and T$\mathbb{Z}$T, and possibly even TST + Infinity from Models of *i*NF

Definition 83
$\mathfrak{P}(x)$ is $\{\sim\sim y : y \subseteq x\}$—or $\{\sim\sim y : y \subseteq \sim\sim x\}$ (which is the same thing).

$\mathfrak{P}$ is for $\mathfrak{P}$owell of course.

This is worth spelling out. Clearly every double complement of a subset of X is a stable subset of the double complement of X. For the inclusion in the other direction reflect that if Y is a stable subset of $\sim\sim X$ then Y is the double complement of $Y \cap X$:

$$\sim\sim(Y \cap X) \;=\; (\sim\sim Y) \cap (\sim\sim X) \;=\; \sim\sim Y = Y.$$

The $\mathfrak{P}$owell operation is a composition of two $\subseteq$-monotone operations and so is $\subseteq$-monotone itself.

You might expect, as i did, that we ought to be able to prove that $\mathfrak{P}(X)$ is stable if X is. We seem to need it. But remember that in $\mathfrak{M}$, the universal set of level $n+1$ is NOT the collection of inhabitants of level n but is its double complement. So, $V_{n+1} = \mathfrak{P}(V_n)$ but that is not the same as the (internal!) universal set at that level. So we don't need it. Just as well, co's it ain't true!

We need the following lemma.

Lemma 84
Let X, A, B be three sets, with $A, B \subseteq X$.
Suppose $(\sim\sim A) \cap X = (\sim\sim B) \cap X$;
then $\sim\sim A = \sim\sim B$.

Proof:
We have

(1)	$\sim\sim A =$	$\sim\sim A \cap \sim\sim X$	because $A \subseteq X$
(2)	$=$	$\sim\sim(\sim\sim A \cap X)$	push the '$\sim\sim$' in to get the RHS of (1)
(3)	$=$	$\sim\sim(\sim\sim B \cap X)$	because $\sim\sim A \cap X = \sim\sim B \cap X$
(4)	$=$	$\sim\sim B \cap \sim\sim X$	from (3) by pushing the $\sim\sim$
(5)	$=$	$\sim\sim B$	from (4) because $B \subseteq X$

■

This is enuff to ensure that

Corollary 85
Every $\mathcal{L}(TST)$-structure $\langle X, \mathfrak{P}(X), \mathfrak{P}^2(X)\ldots\rangle$ is a model of extensionality.

Proof: This is an immediate consequence of lemma 84.

■

Remark 86 $\mathfrak{M} \models$ *Extensionality.*

Proof:

Suppose X and Y at level $n+1$ are coextensive in the sense of the model, so $X \cap V_n = Y \cap V_n$. We want to infer $\neg\neg(X = Y)$. This is where we reach for lemma 84. X and Y are $\sim\sim x$ and $\sim\sim y$ for two subsets x and y of V_n. Level n is a set of stable sets, so lemma 84 applies, and we infer $\sim\sim x = \sim\sim y$, which is to say, $X = Y$. ■

Looks too good to be true.

For me the obvious next question is: can we show that *i*NF cannot refute any instance of the ambiguity scheme **for stable formulæ restricted to the** K_i. It may be that the restriction in boldface turns a putatively impossible challenge into a conjecturally possible one.

Now that we have the operation $\mathfrak{P}$, corollary 85 tells us that the obvious way to get models of TST and T$\mathbb{Z}$T is to take each level to be $\mathfrak{P}$ of the level below it. For TST we take level 0 to be ... well, not sure ... and for T$\mathbb{Z}$T we simply add countably many constants, $K_i : i \in \mathbb{Z}$. For the moment we consider only putative models $\mathfrak{M}$ of classical TST.

What about comprehension? If $\phi(x, \vec{y})$ is a formula in the range of the negative interpretation then $\{x : \phi(x, \vec{y})\}$ is a stable set and if the 'x' ranges over K_i then $\{x : \phi(x, \vec{y})\}$ is a stable subset of K_i and is therefore a member of K_{i+1}.

In the classical setting the universal set V_{n+1} of level $n+1$ is a member of level $n+1$ and is actually the same thing as level n.

Notice that in our setting the universal set V_{n+1} of level $n+1$ is not the same as level n:

Level $n+1$ is $\mathfrak{P}$(level n);
The universal set V_{n+1} of level $n+1$ is $\sim\sim\mathcal{P}(V_n)$.

[Have i got that right?] We need careful hygiene in our notation to keep this in mind and not trip ourselves up.

Is $\mathfrak{P}(X)$ the same as the set of separable subsets of $\sim\sim X$?

Suppose Y is a separable subset of $\sim\sim X$; therefore for any $y \in \sim\sim X$ we have $y \in Y \vee y \notin Y$. Now if $\neg\neg(y \in Y)$ we cetainly have $\neg\neg(y \in X)$ whence $y \in Y \vee y \notin Y$ by separability. $y \notin Y$ is impossible by assumption so we get the other horn. So $Y \in \mathfrak{P}(X)$.

For the other direction suppose $Y \in \mathfrak{P}(X)$. Y is a double complement of a subset y of X. We want

$$(\forall x)(\neg\neg(x \in X) \to (\neg\neg(x \in y) \vee x \notin y))$$

What can we do with $\neg p \to (\neg q \vee \neg r)$?

We can get it from $(\neg p \to \neg q) \vee (\neg p \to \neg r)$ and we can get that from $(q \to p) \vee (r \to p)$. p is $\neg(x \in X)$; q is $\neg(x \in y)$; r is $x \in y$

So the answer appears to be: no!

Notice that we haven't spelled out which negative interpretation is in play, and are using only the fact that every formula in the range of the negative interpretation is stable.

Lemma 78 tells us that notnotequality at any level > 0 is simply equality, and that simplifies things mightily.

And it's the same with the membership relation... between levels it's ordinary membership, and its graph is stable beco's everything to the right of an '$\in$' is stable.

We now have to consider what level 0 is to be. If we take it to be something that is not kfinite (like V) then perhaps the resulting model will be a model of TST + Infinity. Work to do here.

Let's give it some thought. Level 0 of $\mathfrak{M}$ is (say) $V/(\neg\neg =)$. We want $\mathfrak{M} \models V/(\neg\neg =)$ is not kfinite iff $V/(\neg\neg =)$ is in fact not kfinite. $V/(\neg\neg =)$ is not kfinite iff it is NOT the case that every set containing $\emptyset$ and closed under etc etc contains $V/(\neg\neg =)$. $\mathfrak{M}$ is going to believe that $V/(\neg\neg =)$ is not kfinite iff it is NOT the case that every set in $\mathfrak{M}$ containing $\emptyset$ and closed under etc etc contains $V/(\neg\neg =)$. So we want it to be a sufficient condition for $V/(\neg\neg =)$ to be Kfinite that every STABLE set containing $\emptyset$ and closed under etc etc contains $V/(\neg\neg =)$. It would actually be enough for that to be a suff condition for $V/(\neg\neg =)$ to be notnotKfinite but that probably doesn't help.

So we want: if every *stable* set containing $\emptyset$ and closed under etc etc contains $V/(\neg\neg =)$, then *every* set containing $\emptyset$ and closed under etc etc contains $V/(\neg\neg =)$.

We are told that every *stable* set containing $\emptyset$ and closed under etc etc contains $V/(\neg\neg =)$. So sse X is a set containing $\emptyset$ and closed under etc etc. We want to prove that it contains $V/(\neg\neg =)$. How about $\sim\sim X$. It's stable all right. Is it closed under etc etc? If so, it contains $V/(\neg\neg =)$. Then we'd have to show that this is enuff to show that X contained $V/(\neg\neg =)$.

Here the discussion about Horn formulæ in constructive Logic around lemma 4 on p. 10 could help, but it does seem that one needs to be more careful in one's choice of a stable set. The stable sets of concern to us are not just stable, they are members of V_2.

Suppose $\mathfrak{M}$ does not believe that its bottom level is finite. Is this going to propagate upwards? Presumably it does, beco's $\mathfrak{P}$ obeys classical logic. But one might sensibly ask: does $\mathfrak{P}$ preserve kfiniteness? Or Nfiniteness? Presumably not ... How would a proof go? Suppose X and $\mathfrak{P}(X)$ are both kfinite. $\mathfrak{P}(X \cup \{x\})$ is the set of double complements of subsets of $X \cup \{x\}$. We would like to obtain this from $\mathfrak{P}(X)$ and $\mathfrak{P}(\{x\})$ somehow. The first is kfinite but the second isn't.

In this context it seems natural to ask whether or not the set of stable sets is finite. Suppose the set of things that are not double complements is finite. Then the set of double complements of its members (which i *think* is the set of stable sets—I'm betting that every stable set is the double complement of something that is not stable) is also finite. So their union is finite. That union is the set of sets that are stable or not stable. That *ought* to be V but of course it won't be, but it should at least be infinite. That's the idea anyway.

In what sense is $\mathfrak{M}$ transitive? Suppose $y \in x \in V_{n+1}$. Since $x \in V_{n+1}$ it is the double complement of a subset w of V_n. So $\neg\neg(y \in w)$ and $w \subseteq V_n$ so $\neg\neg(y \in V_n)$. That's all one should expect.

Gulp. But isn't V_n stable?

'stable' is 1-symmetric. But it is not preserved under cardinality. $\{\emptyset\}$ is stable but it is the only singleton that is. But what about bigger sets? Is there a simple condition such that two sets with that condition that are bijective copies are either both stable or both not stable?

12.6 Beeson Interprets (classical) T$\mathbb{Z}$T in *i*NF

This is what he says

1. Negative interpretation of classical NF into *i*NF plus a new unary predicate $P(x)$ with the axiom

$$P(y) \leftarrow \forall x(((\forall z(z \subset x \land \mathit{Stable}(z) \to z \in x))))$$

This has become garbled

2, Now suppose NF proves ϕ. Then for some conjunction Γ of axioms of NF, there is a cutfree Gentzen proof of $\Gamma \vdash \phi$.

3. So the formulas in this proof can be simultaneously stratified with depth at most N for some N.

4. We can define in *i*NF a formula P^* that says x is stable hereditarily up to N levels down. So replacing P by P^* we get a proof of the sequent $(\Gamma^-)^* \vdash (\phi^-)^*$. [$\Gamma^-$ is the double negation version of Γ, and $(\phi^-)^*$ is the double-negation version of ϕ using P^* instead of P.]
The starred version of the axiom for P is provable constructively.
But that eliminates P and the axiom for P, starred, is provable in *i*NF, so we get a proof in *i*NF of the above-mentioned sequent.

5. Now take ϕ to be falsity. Then an inconsistency in NF converts to an inconsistency in *i*NF.

Key idea: a *uniformly (or simultaneously) stratified proof*... and a key fact to go with it. Every simultaneously stratified proof in NF of a stratified formula ϕ corresponds in a straightforward way to a proof in TST + the scheme "there are at least n objects at level 0" of the formula obtained from ϕ by incorporating into the variables of ϕ the naturals used in a stratification of ϕ.
There is another thing one can do with uniformly stratified proofs in NF. For a suitable negative interpretation (which Beeson writes with a '-') one can

show that from a uniformly stratified proof $\mathcal{D}$ in NF of ϕ one can obtain a proof $\mathcal{D}'$ of ϕ^- in [a slightly spiced up version of] *i*NF.

This will show that if there is a proof of the false in TST then there is a proof of the false in *i*NF.

So far so good. Now we have to come clean on what the negative interpretation is that Beeson is using (tho' the precise details probably don't matter very much) and we have to explain what we have to do to obtain $\mathcal{D}'$.

The first thing to do is fix a uniform stratification and then attach to each variable the natural number associated to it by the uniform stratification. Then we restrict variables of level n to a predicate $P_n(\ )$ where $P_0(x)$ says $x = x$, and $P_{n+1}(x)$ says that x is the double complement of a set of things all of which are P_n. (In effect this is Powell's definition rather than Beeson's but i think it'll work better). By *restrict* we mean that, to obtain $\mathcal{D}'$ from $\mathcal{D}$, whenever we see '$(\forall x)(\cdots)$' and 'x' has had the natural number n attached to it then we replace it with '$(\forall x)(P_n(x) \to \cdots)$' and analogously for the existential quantifier. We earnestly desire that this should be a proof in *i*NF of something like the negative interpretation of ϕ—or can at least be turned into one.

At this point the ms breaks off.

So let's put that on one side and restart!

12.7 Attempting to interpret NF in *i*NF

Remark 87 *Constructive ZF proves that there can be no set whose double complement is V.*

Proof: Suppose $\sim\sim A = V$. Consider $R = \{x \in A : x \notin x\}$. If R is a member of itself then it isn't, so it isn't. So it doesn't satisfy the membership condition for R, whence $\neg(R \in A \wedge R \notin R)$. However it does satisfy the second conjunct, so it must fail the first, whence $R \notin A$. But this contradicts $\sim\sim A = V$. ■

That used unstratified separation. It looks suspiciously easy (Michael Rathjen had been thinking about the problem of proving in constructive ZF that there is no set whose double complement is the universe) but Michael says it's OK. Makes you wonder about the stratified constructive fragment. Might there be some profit in considering a constructive version of the Baltimore model in [22]?

check extensionality and the rud functions

Lemma 88 *In any theory that has Γ-aussonderung, the assertion that every set is stable implies excluded middle for formulæ in Γ.*

Proof:

Let ϕ be an expression in Γ and think of $\{x \in \{1\} : \phi\}$. By stability this set is either $\{1\}$ or is empty. So ϕ must be true or false. ■

This tells us that the class of hereditarily stable sets must be a model of excluded middle, since **Z** has aussonderung for all formulæ.

We then modify the interpretation * so that $(\forall x)(\phi^*(x))$ is $(\forall x)(\neg(Hstab(x) \wedge \neg\phi^*(x))$ and $(\exists x)(\phi^*(x))$ is $\neg(\forall x)\neg(Hstab(x) \wedge \phi^*(x))$. Both of these are stable because any negated formula is stable.

The clauses for quantifiers involve restriction to a set with nice properties. One can get an interpretation of classical **Z** into intuitionistic **Z** by using the above and restricting all quantifiers to the collection of hereditarily $\neg\neg$stable sets.

The same will work in NF if we can find a set which is equal to the set of its $\neg\neg$stable subsets and is itself $\neg\neg$stable.

The recursive clauses above are designed to ensure that, for all ϕ, $\{y : \phi^*(\vec{x}, y)\}$ is a stable set. (Actually we need it only to be $\neg\neg$stable so this is overkill). That is easy enough to check. Extensionality is slightly different. We want

$$(\forall x \in X)(\forall y \in X)((\forall z \in X)(\neg\neg(z \in x) \longleftrightarrow \neg\neg(z \in y)) \to \neg\neg(x = y)$$

Since X is equal to the set of its $\neg\neg$stable subsets the restriction of the z to X does nothing, so the antecedent merely says that $\sim\sim x = \sim\sim y$. But if x and y are $\neg\neg$stable we have $\neg\neg(x = \sim\sim x)$ and $\neg\neg(y = \sim\sim y)$, whence $\neg\neg(x = y)$ as desired.

Now suppose the greatest fixed point is a set. Call it $\mathcal{S}$. We have

$$x \in \mathcal{S} \longleftrightarrow (\exists z)(z \subseteq \mathcal{P}_{stab}(z) \wedge x \in z)$$

and we want $\mathcal{S} \in \mathcal{S}$, so that it will be a model for *i*NF. To this end we want $\neg\neg(x \in \mathcal{S}) \to x \in \mathcal{S}$. But there seems no reason to expect $\neg\neg(\exists z)(z \subseteq \mathcal{P}_{stab}(z) \wedge x \in z) \to (\exists z)(z \subseteq \mathcal{P}_{stab}(z) \wedge x \in z)$?

How about the greatest fixed point for $\lambda x.\mathcal{P}_{\neg\neg stab}(x)$? We have the same problem: $\neg\neg\exists \not\to \exists\neg\neg$. In both cases the $\neg\neg\exists$ arises because we are looking at a *greatest* fixed point. We'd have a $\neg\neg\forall$ with the *least* fixed point but that's no use to us.

In fact we do not even need X to be equal to the set of its $\neg\neg$stable subsets: it would suffice to have an X that was mapped by a permutation of V onto the set of its $\neg\neg$stable subsets. The only trouble is that there is no reason to suppose there is such a set or such a permutation.

13 *i*NFU + "all atoms are notnot equal"

Randall,

You have convinced me that "any pair

First idea:

Construct a kripke model starting from a model $\mathfrak{M} = \langle M, \in \rangle$ (of, as it might be NFU) where each world has M for its carrier set, and the worlds are indexed by finite sets of atoms. The idea is that each world believes that the atoms in its label are equal, in that they belong to the same thing. So W_A believes $a \in x$ whenever $a, b \in A$ and $b \in x$ (or a is not an atom, and $a \in x$). So we have to omit $=$ from our language, and introduce it as a defined expression. The idea is that eventually any two atoms are notnotequal.

What are these finite sets of atoms exactly? I said above that $W_A \models x \in y$ if $x \in y$ or $x \in A \wedge b \in A \wedge b \in y$. If A appears as a variable then it gets the same level as y in any stratification and this will mean that the assertion that $W \models x \in y \wedge y \in z$ is unstratified. More generally it would mean that "$W_A \models \phi$" is unstratified even if ϕ is stratified, and we don't want that. So we index the possible worlds not by sets of atoms but by metalanguage sets of names of atoms. That is the second idea.

$W_A \models x = y$ iff

$(\forall W' \supseteq A)(\forall z)(W_{A'} \models x \in z$ iff $W_{A'} \models y \in z)$ which is to say

iff $\mathfrak{M} \models$ (neither x nor y contain any atoms and $x = y$) or $\mathfrak{M}$ believes x and y are both atoms and are both in A or....

As part of the project to understand the strength of *i*NF one looks for propositions that have what i like to call *leverage*. I think we have enjoyed sufficiently little success so far for it to be worth while casting our net a little wider and thinking about *i*NFU. Work of Randall's going back to 1998 (and further back, less directly) to Crabbé's work on SF, shows that there is a negative interpretation from NFU into *i*NFU. Old work of Boffa showed that if there are sufficiently few atoms in a model of NFU then one can interpret NF.

There are all sorts of conditions one can postulate for the set of atoms. The time has come to consider whether or not, in this constructive context, they have any *leverage*.

It's obvious that *i*NFU + "notnot all atoms are equal" is strong; what is less obvious is that *i*NFU + *all atoms are notnot equal* is strong too. This may matter.

Work in *i*NFU + *all atoms are notnot equal*.

Suppose V is Kfinite. Since V is finite then the collection of atoms is subfinite so we can exploit Linton-Johnstone to infer not-not all atoms are equal. But "all atoms are equal" implies NF, which implies that V is not Kfinite, contradicting assumption and proving the false. So "notnot (all atoms are equal)" contradicts "V is Kfinite" too. So we have shown:

Remark 89 *iNFU + all atoms are notnotequal* $\vdash$ *V is not Kfinite.* *(A)*

That is to say: *i*NFU + *all atoms are notnotequal* is **strong**.

Perhaps this is not surprising. After all, it's obvious that

$$i\text{NFU} + \text{notnot(all atoms are equal)} \vdash V \text{ is not Kfinite.} \qquad \text{(B)}$$

...beco's it's obvious that *i*NFU + notnot(all atoms are equal) is strong—the classical version is NF, after all. But (A) is slightly stronger than (B)—and less obvious.

The remark has the consequence that there is a bound on what my forcing construction can do. That construction throws up models in which all atoms are notnot equal. Does it ever throw up a model of *i*NFU? By the remark, any model of NFU that it comes up with will believe that V is not kfinite, so it will only throw up models of NFU if the model of NFU that we input into the construction is strong.

It looks pretty discouraging: in order to get a model of *i*NFU out of it *at all* one has to start with a *strong* model of NFU. And i cannot (at this stage) see where in the development of my model one could use the special properties of the strong model that one is using to ensure that the output model satisfies comprehension. I may have bitten off more than i can chew.

One can make the following remarks:

• There is a chink of light, in that any model that my construction gives does at least satisfy the constant domain axiom. I don't know what the constant domain axiom does for one in this NFiste context: no-one has ever looked at it (as far as i know). My guess is that it doesn't do a great deal but—as i say—no-one has ever checked it.

• NB 'strong' mightn't neccessarily mean "implies infinity". It means "high consistency strength" It may be that NFU + not-infinity plus "every wombat is a dingbat" is a strong theory that one can feed into my construction. Are there natural strong axioms one can add to NFU + not-infinity to get things as strong as NFU + infinity? Transitive closure? That sort of thing...

• Mind you, the above musing does not imply that *i*NFU + *all atoms are notnotequal* is as strong as NF, merely that it is as strong as *i*NFU + infinity ...which prompts the question...

• Do we know of any models of *i*NFU + *all atoms are not-not equal*? Well, yes we do, beco's *grace à toi* we know that NF is consistent. But is there a more idiomatic constructive NFU-ish proof? Do we have any other methods that will produce actual models of this theory?

• Can we interpret NFU + Infinity in *i*NFU + all atoms are not-not equal? A nice negative interpretation...?

Actually Rndall has pointed out something the scuppers this programme completely. Let a and b be distinct atoms, and consider the fates of $\{a\}$ and $\{b\}$ in the Kripke structure. They are distinct, beco's $\{\{a\}\}$ contains $\{a\}$ but not $\{b\}$. However, anything notnot in one is notnot in the other: $(\forall x)(\neg\neg(x \in \{a\}) \longleftrightarrow \neg\neg(x \in \{b\})$ and $(\forall x)\neg\neg(x \in \{a\} \longleftrightarrow x \in \{b\})$. But then by

Linton-Johnstone we can pull out the $\neg\neg$ and infer that $\{a\}$ and $\{b\}$ are notnot coextensive. So they ought to be notnotequal. But they are unequal.

Bugger. So weak extensionality is false. Not forall [stuff]. But we haven't actully obtained a counterexample!

Let T be the constructive theory of the Kripke model. Write '$A(x)$' for "x is a pair of nonempty coextensive things", and '$B(x)$' for "they are equal". Randall has convinced me that the model satisfies $(\exists x)(\neg\neg A(x) \wedge \neg B(x))$. And that is enuff to prove $\neg(\forall x)(A(x) \to B(x))$.

So far so good. But does T actually prove the existence of a counterexample? Observe, as i observed, that there doesn't seem to be a constructive proof of the sequent

$$\neg(\forall x)(A(x) \to B(x)),\ \neg(\exists x)(A(x) \wedge \neg B(x)) \vdash$$

Does $T \vdash (\exists x)(A(x) \wedge \neg B(x))$? It might, but not simply beco's it proves $\neg(\forall x)(A(x) \to B(x))$ So suppose T *doesn't* prove there is a counterexample. Then we can consistently add an axiom to say that we cannot find a-with-b which are at once coextensive, nonempty *and* distinct, that is to say

$$\neg(\exists x)(A(x) \wedge \neg B(x)).$$

Now this formula constructively implies

$$(\forall x)(A(x) \to \neg\neg B(x))$$

This is why i want an *actual* counterexample. If the theory doesn't produce one, then we can adopt $(\forall x)(A(x) \to \neg\neg B(x))$ as an axiom, and tweak our definition of equality to obtain a model of weak extensionality.

While we are on this subject, here is a further thought. Suppose NF is consistent, so *i*NF is consistent too. So *i*NF has extensions that do not contradict classical logic; in fact it has *strong* extensions that do not contradict classical logic. But it has interesting extensions that do contradict classical logic and may be consistent—"there is a finite set whose double complement is V" for example. It's just occurred to me that *i*NF might have extensions that contradict classical logic but are *strong*.

Here is an obvious thing to try. Yet another attempt to get a sensible Kripke model (a silk purse) out of a model of NFU (sow's ear).

Sens**I**b**L**e **K** ri**P**ke model out of a model for NF**U**... **RSE**

Start with a structure $\mathfrak{M} = \langle M, =, \in\rangle$; the construction we are about to unfold was conceived as something one would do to a model of NFU, but $\mathfrak{M}$ could be anything. Expand $\mathfrak{M}$ by adding names for all the atoms, getting a structure in an expanded language which we will call $\mathcal{L}(\mathfrak{M})$. We set up a Kripke model $\mathfrak{K}$ as follows: the language $\mathcal{L}(\mathfrak{K})$ contains $\in$ only—no equality! We will subsequently define $x = y$ to be $(\forall z)(x \in z \longleftrightarrow y \in z)$. This means that substitutivity of equality will have to be *proved*! We will abuse notation to the extent of denoting the expanded structure, too, by '$\mathfrak{M}$'.

For each (externally) finite set A of names, we have a possible world W_A. The domain of W_A—any A—is M. Thus our model will obey the logic of constant domains.

Naturally the accessibility relation on worlds is the inclusion relation on the subscripts.

We will have the usual recursions for constructive Kripke semantics, namely

Definition 90

$$
\begin{array}{lll}
W_A \models x \in y & \textit{is} & \bigvee_{a,b \in A}[(x = b \wedge a \in y) \vee (\exists w)(w \in x \in y)]; \\
W_A \models (\phi \vee \psi) & \textit{is} & W_A \models \phi \vee W_A \models \psi; \\
W_A \models (\phi \wedge \psi) & \textit{is} & W_A \models \phi \wedge W_A \models \psi; \\
W_A \models (\phi \to \psi)^A & \textit{is} & (\forall A' \supseteq A)(W_{A'} \models \phi \to W_{A'} \models \psi) \\
W_A \models ((\forall x)\phi(x)) & \textit{is} & (\forall A' \supseteq A)(\forall x \in W_{A'})(W_{A'} \models \phi(x)).
\end{array}
$$

$\neg\phi$ is $\phi \to \bot$ as usual, so $W_A \models (\neg\phi)$ is $(\forall A' \supseteq A)(W'\neg \models \phi)$

However, since all our possible worlds have the same domain, we can simplify

$$W_A \models (\forall x)\phi(x) \text{ iff } (\forall A' \supseteq A)(\forall x \in W_{A'})(W_{A'} \models \phi(x))$$

to

$$W_A \models (\forall x)\phi(x) \text{ iff } (\forall x)(W_A \models \phi(x))$$

We'd better say a little bit about why this is so.

Suppose $(\forall x)(W_A \models \phi(x))$. Then, by persistence, we infer $(\forall x)(W_{A'} \models \phi(x))$ for any $W' \supseteq W$. Now, since all the worlds have the same carrier we, we can infer $(\forall A' \supseteq A)(\forall x \in W_{A'}(W_A \models \phi(x))$ which is the definition of $W_A \models (\forall x)(\phi(x))$.

That completes the declaration of the model.

(The reader might be wondering: why do we have names for atoms, and not just variables ranging over finite sets of atoms? The point is that, were we to go down that road, translations of $x \in y \wedge y \in z$ would have stratification conflicts at the site of the variable ranging over the sets of atoms.)

The question now is: what does this model actually satisfy?

There are two things one might mean by this. With Kripke models i can never remember whether the model believes something iff all worlds believe it, or believes it iff the designated world believes it. We don't need to decide which, beco's fortunately this model $\mathfrak{K}$ exhibits a phenomenon i like to call *persistence*:

Remark 91 *If $X \subseteq Y$ and $W_X \models \phi$ then $W_Y \models \phi$.*

Proof:

We prove this helpful and comforting fact by recursion on the subformula relation. Clearly true for atomics; the induction steps for propositional connectives are easy.

∃:

Suppose $X \subseteq Y$ and $W_X \models \exists x\phi(x)$. Then there is a witness, $a \in \text{dom}(W_X)$, s.t $W_X \models \phi(a)$. But all worlds have the same domain, so this a also inhabits W_y and, by induction hypothesis, we have $W_Y \models \phi(a)$, whence $W_Y \models \exists x\phi(x)$.

∀:

Suppose $X \subseteq Y$ and $W_X \models \forall x\phi(x)$. That is to say, for all $X' \supseteq X$ and all $a \in \text{dom}(W_{X'})$, $W_{X'} \models \phi(a)$. For W_Y to believe $\forall x\phi(x)$ it is necc and suff that for all $Y' \supseteq Y$ and all $a \in \text{dom}(W_{Y'})$, $W_{X'} \models \phi(a)$. But this follows from $Y \supseteq X$.

That seems to cover all cases.

■

We are going to be interested in the set of all things believed by the root world $W_\emptyset$ ("forced by the empty condition"). Let us call this theory 'T' for the moment. It would be nice to know what T is. It would be nice if T obeyed substitutivity of identity, and i hope to prove that it does. Some things are clear.

(i) Since all the worlds in the model have the same domain/carrier-set, T contains the constant domain axiom CD:

$$(\forall x)(A(x) \vee B) \to .(\forall x)(A(x)) \vee B \qquad \text{CD}$$

'x' not free in B of course.

(ii) T contains $(\forall \text{ atoms } a, b)(\neg\neg(a = b))$.

(i) doesn't seem to do very much for us, but (ii) turns out to be strong, and in a way that will obstruct proof of the comprehension axioms.

Remark 92 *iNFU + all atoms are notnotequal* $\vdash V$ *is not Kfinite.*

Proof:

Work in *i*NFU + *all atoms are notnot equal.*

Suppose V is Kfinite. Since V is finite then the collection of atoms is subfinite so we can exploit Linton-Johnstone to infer not-not(all atoms are equal). But *i*NFU + "all atoms are equal" is *i*NF, which proves that V is not Kfinite, So if V is Kfinite then it isn't Kfinite. So it isn't Kfinite.

■

This will mean that if we execute this construction starting with a model $\mathfrak{M}$ of NFU that does *not* satisfy "V is infinite" then the $\mathfrak{K}$ that we get out of it will not be a model of *i*NFU, beco's if it were it would have to be a model of *i*NFU + infinity, and we can't prove the consistency of NFU + infinity relative to the consistency of NFU + ¬Infinity.

Let's prove (ii)

$\mathfrak{K} \models$ "any two atoms a and b are notnotequal" because all sufficiently late conditions contain both names, and therefore believe that any set containing one also contains the other.

■

With a view to nailing down the theory T, we define, by recursion on formulæ, a map $(\phi, A) \mapsto \phi^A$ that takes a formula $\phi \in \mathcal{L}(\in)$ and a finite set A of names of atoms, and gives a formula in $\mathcal{L}(\in, =)$. The intention is that $W_A \models \phi$ iff $\mathfrak{M} \models \phi^A$. There is also the intention that ϕ^A shall be stratified if ϕ is.

What we want to say (the effect we want to achieve) is a recursive definition of $()^A$ which echoes the recursive semantics for possible worlds. $\perp^A$ is $\perp$ of course; there is only one atomic case:

DEFINITION 93

$(x \in y)^A$ *is* $\bigvee_{a,b \in A}[(x = b \wedge a \in y) \vee (\exists w)(w \in x \in y)]$;
$(\phi \wedge \psi)^A$ *is* $\phi^A \wedge \psi^A$;
$(\phi \vee \psi)^A$ *is* $\phi^A \vee \psi^A$;
$(\phi \to \psi)^A$ *is* $\bigwedge_{A' \supseteq A}(\phi^{A'} \to \psi^{A'})$
(so $(\neg\phi)^A$ *is* $\bigwedge_{A' \supseteq A}(\neg(\phi^{A'}))$*);*
$((\forall x)\phi)^A$ *is* $\bigwedge_{A' \supseteq A}(\forall x)(\phi^{A'})$.

But we can't say that *outright* because $\bigwedge_{A' \supseteq A}$[stuff] involves an infinite conjunction, every finite set A of names having infinitely many supersets A'. This is not to be borne! The purpose of the following result is to ensure that ϕ^A is genuinely a finite string, and actually has some very nice properties. For example, the theory forced by the empty condition doesn't explicitly mention any atoms.

`\begin{challenge}\mbox{\negthinspace}\\` For all ϕ in some class Γ of formulæ yet to be delineated, ϕ^A is always equivalent to an expression of $\mathcal{L}(\mathfrak{M})$ mentioning only names in A.

I think one such Γ is formulæ in prenex normal form, or rather a slightly expanded class consisting of formulæ with all quantifiers exported (pulled to the front) followed by a formula that is a conjunction of disjunctions of conjunctions of ... atomics and negatomics and negnegatomics.

IN fact the class (Let us call it the class of translatable formulæ) we are interested in is the class of atomics, negatomics and negnegatomics closed under $\wedge$, $\vee$ and the two quantifiers. It's only the clause for $\to$ that requires us to look at arbitrary finite extensions of conditions, and sabotages the recursion. Negation happens to work on atomics. Crucially every formula is classically equivalent to a formula in this class. So the fact that we can add to T anything classically equivalent to something in T means that we can get the whole of *i*NFU.

So Con(NFU) implies Con(*i*NFU + $(A \to B) \vee (B \to A)$ plus CD plus any two atoms are notnotequal.

The point is that this $\{-\}^A$ map gives rise to a relative consistency proof. The set of formulæ ϕ believed by $W_\emptyset$ s.t. $\mathfrak{M} \models \phi^\emptyset$ is a theory consistent relative to $Th(\mathfrak{M})$.

We can define ϕ^A for atomics, negatomics and negnegatomics, and the recursions for both quantifiers and $\vee$ and $\wedge$ (but not $\to$) work. $\to$ is the only one that causes an infinite conjunction.

Let's start by spelling out ϕ^A for atomics, negatomics and negnegatomics.

$x \in y$, $x \notin y$, $\neg\neg(x \in y)$

(For the moment we don't worry about $x = y$, $x \neq y$, $\neg\neg(x = y)$ co's $\mathcal{L}(\mathfrak{M})$ doesn't contain '=')

13.1 $(x \in y)^A$

This is

$$(x \in y) \vee \bigvee_{a,b \in A} (x = b \wedge a \in y)$$

from which we can export the $\bigvee$:

$$\bigvee_{a,b \in A} [(x \in y) \vee (x = b \wedge a \in y)]$$

This certainly has the desired form: it is an expression of $\mathcal{L}(\mathfrak{M})$ which doesn't use any names not in A.

13.2 $(x \notin y)^A$

This is

$$(\forall A' \supseteq A) \bigwedge_{a,b \in A'} \neg[(x \in y) \vee (x = b \wedge a \in y)]$$

$$(\forall \text{ atoms } a, b)[(x \notin y) \wedge (x \neq b \vee a \notin y)]$$
$$(x \notin y) \wedge (\forall \text{ atoms } a, b)[x \neq b \vee a \notin y]$$

which doesn't mention A.

Some duplication here!

How about $(x \notin y)^A$? It must be

$$\bigwedge_{A' \supseteq A} [(x \notin y) \wedge \bigwedge_{a,b \in A'} (x \neq b \vee a \notin y)]$$

Now we can export the $\bigwedge_{a,b \in A'}$ to get

$$\bigwedge_{A' \supseteq A} \bigwedge_{a,b \in A'} [(x \in y) \wedge (x \neq b \vee a \notin y)]$$

which simplifies to

$$(\forall u, v)[\text{empty}(u) \wedge \text{ empty}(v) \to (x \notin y) \wedge (x \neq v \vee u \notin y)]$$

'atom(x)' is of course short for '$(\forall z)(z \notin x)$'.

13.3 $\neg\neg(x \in y)^A$

It must be

$$(x \in y) \vee (\mathrm{atom}(x) \wedge (\exists a)(\mathrm{atom}(a) \wedge a \in y)$$

and the 'A' has disappeared.
Now for the recursions:
$((x \in y) \wedge (y \in z))^A$
is
$(x \in y)^A \wedge (y \in z)^A$

$$\bigvee_{a,b \in A} [(\exists w)(w \in x \in y) \vee (x = b \wedge a \in y)] \wedge \bigvee_{c,d \in A} [(\exists w)(w \in y \in z) \vee (y = d \wedge d \in z)]$$

which is going to resolve into a horrible thing

$$\bigvee_{a,b,c,d \in A} [((\exists w)(w \in x \in y) \vee (x = b \wedge a \in y)) \wedge ((\exists w)(w \in y \in z) \vee (y = d \wedge d \in z))]$$

[i hope i've calculated that correctly!] which is at least of the right form.
$((x \in y) \vee (y \in z))^A$ is slightly easier

$$\bigvee_{a,b \in A} [(\exists w)(w \in x \in y) \vee (x = b \wedge a \in y)] \vee \bigvee_{c,d \in A} [(\exists w)(w \in y \in z) \vee (y = d \wedge d \in z)]$$

which becomes (i think!)

$$\bigvee_{a,b,c,d \in A} [((\exists w)(w \in x \in y) \vee (x = b \wedge a \in y)) \vee ((\exists w)(w \in y \in z) \vee (y = d \wedge d \in z))]$$

The moral seems to be that if ϕ^A and ψ^A are of any of the forms
$\bigwedge_{a \ldots \in A}$[stuff],
$\bigwedge_{a \ldots \in A} \bigvee_{b \ldots \in A}$[stuff],
$\bigvee_{a \ldots \in A} \bigwedge_{b \ldots \in A}$[stuff] or
$\bigvee_{a \ldots \in A}$[stuff],
then so are all of (ϕ [`connective`] ψ)^A as long as [`connective`] doesn't involve calls to $A' \supseteq A$.

That leaves $\rightarrow$ and $\forall$. However $\forall$ is all right because the constant domain axiom means that the recursion for $\forall$ does not involve a call to $\forall A' \supseteq A$.

So what do we get out of this? It turns out that ϕ^A is defined for atomics, negatomics and negnegatomics, and for formulæ obtained from them by $\wedge$, $\vee$, $\exists$ and $\forall$.

A few things to note:

(i) These rearrangements preserve stratification, so ϕ^A is stratifed as long as ϕ is;
(ii) The set of formulæ reached by this recursion is "dense": every formula is classically equivalent to a formula of our form.
(iii) If ϕ and ψ are **closed** formulæ of the right kind then $(\phi \to \psi)^A$ is a **scheme** $\phi^{A'} \to \psi^{A'}$ over all $A' \supseteq A$. Not much help probably, co's you can only use it *once*!
(iv) We have the constant domain axiom, which might allow us to play around with scopes of quantifiers. But we probably al;ready have all the freedom of manœuvre we need.

Now! What was the point of this ϕ^A caper? I think the point is that if $\phi \in T$ and $\phi^\emptyset$ is defined, then $\mathfrak{M} \models \phi^\emptyset$. In other words, if $\mathfrak{M}$ is a model of a theory T^* then this construction gives us an interpretation of T into T^*. So, it seems, we have a construction that takes a model of a classical theory T^* an emits a model of a constructive theory that satisfies CD plus all atoms are notnot-equal plus, for every axiom of T^*, a formula constructively equivalent to it.

$W_\emptyset \models x = y$ iff, for all z, either z contains no atoms in which case $x \in z \longleftrightarrow y \in z$ or z contains some atoms, and both x and y are atoms.

This means that, in the definition of ϕ^A we have to replace clauses like

$$(\phi \to \psi)^A \text{ is } \bigwedge_{A' \supseteq A} (\phi^{A'} \to \psi^{A'})$$

By some instruction like the following: "Take all the expressions of the form $\phi^{A'} \to \psi^{A'}$ with $A' \supseteq A$ and find some expression in the language-of-set-theory-plus-names-in-A which has the same truth-conditions as the conjunction $\bigwedge_{A' \supseteq A}(\phi^{A'} \to \psi^{A'})$. That will be $(\phi \to \psi)^A$."

We will also want to ensure, as we go along, that ϕ^A is stratifiable as long as ϕ is. If we don't do that then we can wave goodbye to $\mathfrak{K} \models i$NFU.

Proof:

I haven't proved this yet!

■

Obviously this is going to be proved by induction on formulæ.

To power this induction we will need a lemma that says that any formula of the form ϕ^A can be expressed with the $\bigwedge$s and $\bigvee$s exported. Observe that $(x \in y)^A$ starts off with $\bigvee_{a,b \in A}$. We want to export this $\bigvee$ (or $\bigwedge$ if we have negated anything) so that ϕ^A has become $\bigvee_{a,b \ldots \in A} \bigwedge_{a,b \ldots \in A}$[stuff] where [stuff] lies entirely in the language of set theory. (Distributivity means we can have the $\bigwedge$ and the $\bigvee$ the other way round if we prefer.)

The reason for this is that when we put a $\bigwedge_{A' \supset A}$ on the front (as the recursion requires of us) to get one of the two prefixes:

$\bigwedge_{A' \supseteq A} \bigwedge_{a,b \in A'} \bigvee_{a,b \in A'}$[stuff] or

$\bigwedge_{A' \supseteq A} \bigvee_{a,b \in A'} \bigwedge_{a,b \in A'}$[stuff] we can simplify.

The first prefix disappears in favour of a universal quantifier (or several universal quantifiers) over atoms.

This is where danger lurks. We need to take serious;y the possibility that when we

A critical step is the replacement of the $\bigwedge$s and $\bigvee$s by quantifiers, and the replacement of names of atoms by variables. The danger is that thereby we destroy stratification. This matters because it will obstruct the project to make the comprehension axioms true. work to do here

And $\bigwedge_{A'\supseteq A}\bigvee_{a,b,c\ldots inA'}\bigwedge_{a,b\in A'}$ simplifies to $\bigvee_{a,b,c\ldots inA}\bigwedge_{a,b\in A}$ which is of the desired form (the $\bigvee$ and $\bigwedge$ remain exported). (It's just as well that we can do this, beco's the $\bigwedge_{A'\supseteq A}$ is an infinite conjunction!)

So: what we do we do? We prove by induction on the subformula relation that ϕ^A is equivalent to something which starts off with a $\bigvee_{a,b\ldots\in A}$. We have a clause for each clause in definition 93.

OK: we prove by induction on formulæ that every ϕ^A is equivalent to a formula wherein all the reference to atoms are bundled together in a $\bigwedge_{a,b,c\ldots inA}$ or a $\bigvee_{a,b,c\ldots inA}$ at the front of the formula. When we put a $\bigwedge_{A'\supseteq A}$ outside they cancel out. This needs to be spelt out

Since, in $\mathfrak{K}$, we take equality to be the relation of belonging-to-the-same-things, this has the effect that W_A believes that any two atoms with names in A are equal.

We adopt the usual semantics. We have to establish that, for every $\phi \in \mathcal{L}(\in)$, there is $\phi^A \in \mathcal{L}(\in,=)$ satisfying $\mathfrak{K} \models \phi$ iff $\mathfrak{M} \models \phi^A$.

I think we have to show that, for every ϕ and every finite A, there is ϕ^A which mentions only the names in A, and s.t. $W_A \models \phi$ is equivalent to $\mathfrak{M} \models \phi^A$. Certainly $\bot^A$ is $\bot$!

I think ϕ^A is obtained from ϕ simply by replacing '$x \in y$' by the quoted formulation above, at least when ϕ is atomic.

OK, so how does the induction work for $\phi \to \psi$?

What is it a model *of?*

The hope is that $\mathfrak{K}$ is a Kripke model for *i*NFU +

$$(\forall x, y)(\neg(\exists w)(w \in x) \land \neg(\exists w)(w \in y) \to \neg\neg(\forall z)(x \in z \longleftrightarrow y \in z)). \qquad (1)$$

(Any two empty sets are notnot equal.) We noted above that $W_X \models a = b$ whenever $a, b \in X$.

If it works it's fun, but it doesn't give a consistency proof of *i*NF because the displayed formula (1) above (which is equivalent to $(\forall x, y)\neg\neg(\neg(\exists w)(w \in x \lor w \in y) \to x = y)$ has the '$\neg\neg$' the wrong side of the '$\forall$'.

As remarked above we define $x = y$ as $(\forall z)(x \in z \longleftrightarrow y \in z)$. We have to justify substitutivity of $=$.

Substitutivity of Equality

Reason in some world W_X. Suppose $x = y$, which is to say $(\forall z)(x \in z \longleftrightarrow y \in z)$. We aspire to prove $\phi(x) \to \phi(y)$, for all ϕ. This we do by induction on formulæ.

Base case:

Two base cases: $\phi(x)$ could be $x \in w$ or $w \in x$.

The first one is easy.

As for the second: $w \in x$ and $(\forall z)(x \in z \longleftrightarrow y \in z)$. We wish to infer $w \in y$. Specialise 'z' to '$B(w)$'; that should do it.

The induction steps involve quantifiers and connectives.

No problem with $\wedge$ and $\vee$. How about $\to$?

Suppose $W_X \models \phi(x) \to \psi(x)$ and $x = y$.

That is to say

$$(\forall X' \supseteq X)(W_{X'} \models \phi(x) \Rightarrow W_{X'} \models \psi(x))$$

and

$$(\forall z)(x \in z \longleftrightarrow y \in z)$$

We desire to show

$$(\forall X' \supseteq X)(W_{X'} \models \phi(y) \Rightarrow W_{X'} \models \psi(y))$$

So let $X' \supseteq X$ be arbitrary and assume $W_{X'} \models \phi(y)$. Then $W_{X'} \models \phi(x)$ by induction hypothesis. Then $W_{X'} \models \phi(x)$ whence $W_{X'} \models \phi(y)$ by induction hypothesis.

That looked suspiciously easy! Still, we have the quantifiers to come. Do the quantifiers

Extensionality

Extensionality looks like a problem too:

$$(\forall x, y)(((\exists w)(w \in x \vee w \in y) \wedge (\forall z)(z \in x \longleftrightarrow z \in y)) \to (\forall w)(x \in w \longleftrightarrow y \in w))$$

"Two coextensive sets with the same members belong to the same things." Suppose a and b have the same members according to W. Then they must have the same members according to any $W' > W$. This is beco's, altho' a set can acquire new members if W' forcibly identifies one of its members with an atom, the fact remains that any new member acquired by a in this way is also acquired by b—in the same way. So all worlds agree on whether or not a and b have the same members. So $W_\emptyset$ thinks they have the same members, and therefore belong to the same things. But then every subsequent W' believes that they belong to the same things ... beco's you can never leave a set once you've joined it. It's like International Socialists[9]. So, yes, it seems to satisfy

$$(\forall x, y)(\neg(\exists w)(w \in x \vee w \in y) \to \neg\neg(x = y)).$$

[9] Except IS can expel you!

Let $a \neq b$ be two atoms, and $X = \{a, \emptyset\}$ and $Y = \{b, \emptyset\}$. Notice that $X \in B(a)$ (but not $B(b)$) and $Y \in B(b)$ (but not $B(a)$). This can never change, so $X \neq Y$. However they have the same double-complement, in that both eventually contain all old atoms.

Miscellaneous

Suppose the set of atoms is subfinite. Then, by Linton-Johnstone, we can export the $\neg\neg$ and have notnot (all atoms are equal) which gives *i*NF.

OK, so we have a Kripke construction which starts from a model (any model) of NFU and gives back a structure which i am pretty sure is a model of iNFU. It is certainly a model of $\forall atoms a, b \neg\neg(a = b)$. The notnot is to the RIGHT of the forall not the LEFT unfortunately. If it were to the left then we could consistently drop the notnot. Now there is this cute fact due to Linton-and-Johnstone that says that if the domain of the $\forall$ is subfinite, then notnot and forall do commute after all.

So the cunning plan is as follows: Start with a model of NFU in which there are only finitely many atoms. Do my Kripke construction to get a model of iNFU plus all atoms are not-not equal. With any luck there will still be only finitely many atoms, so by Linton-Johnstone we have notnot (all atoms are equal). Then of course we can consistently drop the negneg. That is to say, add the axiom that all atoms are equal. I think we can get a negative interpretation of this extended theory into the theory of the Kripke model.

whence it is consistent that all there is only one atom, so we have iNF. This doesn't give an interpretation of iNF in NFU but it gives a relative consistency result. Perhaps by insinuating another negative interpretation (after the invocation of Linton-Johnstone) we can actually get an interpretation.

To get this to work i need to verify that my Kripke construction does what i hope (and think) it does. I also need to know about the cardinality of the set of atoms in models of NFU.

So: please tell me what you know about the sizes of the set of atoms in NFU.

P.S.: I am locked *down* not locked *up*!!

Here is something that worries me a bit. If $\mathfrak{K}$ is to be a model of *i*NFU then $\{x : \phi(x, \vec{y})\}$ is going to have to exist for all weakly stratified $\phi(\vec{y})$. So what is the witness to this set existence assertion? Well, pretty obviously it is going to be the thing that $\mathfrak{M}$ believes to be the witness. But $\mathfrak{M}$ cannot tell the difference between $\{x : \phi(x)\}$ and $\{x : \neg\neg\phi(x)\}$. So it looks as if the Kripke model is going to believe $\neg\neg p \to p$ as long as p is weakly stratifiable. That seems a bit strong .... Doesn't remark 64 say that that is as strong as NF? But perhaps it's OK, beco's remark 64 was in a context where we were assuming full extensionality. Must check to see whether part 6 implies double negation for weakly stratified formulæ really uses full extensionality.

Actually it's not that bad. $\mathfrak{M}$ can tell the difference between x and $\sim\sim x$, at least if x has an atom as a member; $\sim\sim x$ is $x \cup$ the set of *urelemente*

But it remains that we want to show that, for quite a lot of ϕ, $\mathfrak{K}$ believes ϕ iff $\mathfrak{M} \models \phi$.

Indiscernible Atoms

There is an old question in NFU studies about indiscernible atoms. NFU proves that atoms are indiscernible for stratified formulæ. Is indiscernibility for unstratified expression as strong as Con(NF)?

Here the novel question is: what happens constructively?

This is forcing me to think about what we mean by indiscernibility.

In the usual setting (Ramsey, followed by Erdős and Rado, and subsequent Set Theory) we have a total order on the domain and the claim is that all increasing n-tuples satisfy the same formulæ. It is *strictly* increasing n-tuples that is meant, so that the n-tuples are tuples of distinct things. I don't think we are allowed parameters (see below).

However in the NFU (and iNFU) setting there is no suggestion that the set U of *urelemente* is equipped with a wellordering. (If we are working in NFU (or iNFU) plus $\neg$Inf then U can be given a total order of course)

In the second edition of [20] it is proved in NFU that, for each concrete n, whenever t_1 and t_2 are n-tuples of distinct atoms and ϕ is weakly stratifiable with n free variables then $\phi(t_1) \longleftrightarrow \phi(t_2)$. Notice that ϕ is not allowed to have parameters, lest we take ϕ to be $x = a$, where a is an atom. Then indiscernibility would compel all atoms to be identical, which is a bridge too far. It does assume that all the n atoms are distinct. Specifically the proof requires that the two n-tuples be 1-equivalent as sets, i.e., that there is a permutation of V that translates one onto the other.

Anyway that is one thing that indiscernibility might mean: no global total order, no parameters, and the tuples are tuples of distinct things.

We have to think about kind of indiscernibility-for-atoms holds in our model here. If all atoms are notnot-equal then indiscernibility-of-atoms holds vacuously.

We want to say: if a and b are atoms then $\phi(a) \longleftrightarrow \phi(b)$. But that's not true if $\phi(x)$ is $x = a$. So, as elsewhere, we outlaw parameters.

So how about we say: if a and b are *distinct* atoms then $\phi(a) \longleftrightarrow \phi(b)$? But then if all atoms are notnotequal then trivially they are indiscernible in this sense. So we have to drop the distinctness condition. . . .

"if a and b are atoms then $\phi(a) \longleftrightarrow \phi(b)$." And analogously for tuples. What conditions do we have to put on ϕ? Clearly if ϕ is a stable formula then this follows from substitutivity of equality. Are stratfiable formulæ stable?

What happens if we start with a model of NFU in which the atoms are indiscernible?

14 Holmes interprets NFU in *i*NF

This section was written by Holmes and is included here with his permission.

I now find myself wondering whether this construction can be refined into one that interprets NFU into *i*NFU.

We claim that classical *NFU* can be interpreted in intuitionistic *NF*: (*i*NF).

equality: $x =_{new} y$ is defined as $(\forall z)(\neg\neg(z \in x) \longleftrightarrow \neg\neg(z \in y))$.

(tf sez: this is $\sim\sim x = \sim\sim y$)

sethood: $S(x)$ is defined as $(\forall z)(\forall y \in x)(z =_{new} y \rightarrow \neg\neg z \in x))$.

membership: $x \in_{new} y$ is defined as $S(y) \wedge \neg\neg x \in y$.

This mimics the Crabbé collapse of *SF* to *NFU* seen in [13]. The idea is to replace the constructive relation $x \in y$ with the classical relation $\neg\neg x \in y$. When this is done, failures of extensionality occur. The new equality relation corrects for these failures of extensionality. We are then only interested in sets that respect the new equality relation (thus the new sethood relation). Membership is the intended relation, excluding membership in urelements.

Set definitions built from these notions using the usual translations of classical formulas into intuitionistic formulas (via double negation) will work in *i*NF; the translation of comprehension for classical *NFU* succeeds. That the translation of extensionality holds is obvious from the definition of equality.

The general reason that this works is that all sentences built from the predicates defined above using only $\forall$, $\wedge$, and $\rightarrow$ are equivalent to their own double negations. Thus, we are in the world of classical logic as long as we restrict ourselves to these predicates and logical operations. Comprehension in the interpreted theory works because $\exists$ and $\vee$ are replaced with their classical analogues constructed with the permitted connectives; all the interpreted comprehensions are instances of the more general comprehension of *i*NF.

We prove a lemma (for our own consumption, mostly) to verify this:

Definition: We call a formula *classical* if it is built up from doubly negated atomic formulas by the operations $\forall$, $\wedge$, and $\rightarrow$.

Lemma: For each classical formula ϕ, $\phi \longleftrightarrow \neg\neg\phi$.

Proof of Lemma: By structural induction. The atomic case is obvious.

We claim that $\phi \longleftrightarrow \neg\neg\phi$ implies $(\forall.\phi) \longleftrightarrow \neg\neg(\forall.\phi)$. We only need to show $\neg\neg(\forall.\phi) \rightarrow (\forall.\phi)$. From $\neg\neg(\forall.\phi)$ we can deduce $\neg\neg\phi$, from which we can deduce ϕ by hypothesis, and so deduce $(\forall.\phi)$.

We claim that $\phi_i \longleftrightarrow \neg\neg\phi_i$ for $i = 1, 2$ implies $\neg\neg(\phi_1 \wedge \phi_2) \longleftrightarrow (\phi_1 \wedge \phi_2)$. From $\neg\neg(\phi_1 \wedge \phi_2)$ we can deduce $\neg\neg\phi_i$ for both values of i, from which we can deduce ϕ_i for both values of i by hypothesis, from which we obtain $(\phi_1 \wedge \phi_2)$.

We claim that $\phi_i \longleftrightarrow \neg\neg\phi_i$ for $i = 1, 2$ implies $\neg\neg(\phi_1 \rightarrow \phi_2) \longleftrightarrow (\phi_1 \rightarrow \phi_2)$. From $\neg\neg(\phi_1 \rightarrow \phi_2)$, ϕ_1 and $\neg\phi_2$ we can deduce absurdity (because the latter two hypotheses imply $\neg(\phi_1 \rightarrow \phi_2)$). Thus we have shown $\phi_1 \rightarrow \neg\neg\phi_2$, from which by hypothesis we can show $\phi_1 \rightarrow \phi_2$.

The proof is complete.

Using the Lemma, we verify the interpretation as follows. Any formula built up from the predicates of the purported interpretation of classical *NFU* using connectives permitted in classical formulas is itself classical, and so equivalent to its double negation. From this it follows that classical reasoning is permitted as long as we restrict ourselves to such formulas. The fact that the Crabbé collapse works follows using classical reasoning (replacing the predicates $\in$ and $=$ with their double complements); the fact that comprehension for classical *NFU* works is obvious.

From this it follows that any model of *i*NF in which one cannot produce $x \neq y$ such that $(\forall z. \neg\neg z \in x \longleftrightarrow \neg\neg z \in y)$ supports an interpretation of classical *NF*; for the interpretation of classical *NFU* in such a model will find no urelements.

Thus any weak version of *i*NF must contain unequal sets with the same double complements.

Note that Thomas's complement of Boffa object interpretation already established that we can interpret classical *NF* if having equal double complements is equivalent to being not not equal.

14.1 tf on Holmes's interpretation of NFU

Moved to have a look at this twenty years later!

And again 22 years later!! After 22 years i think the question is: can we interpret NFU not just in *i*NF but in *i*NFU? Answering this will require us to look *very* closely at Randall's 1998 construction.

I'm hoping that this will give a negative interpretation of NFU into *i*NFU!

Let us define an interpretation from NFU into *i*NF. The interpretation will send $=$ to $=_{NFU}$ and $\in$ to $\in_{NFU}$, which we will procede to define. Then there are recursive steps for the quantifiers.

Definition 94 *We define*

$x =_{NFU} y$ *to be* $\sim\sim x = \sim\sim y$;
$\texttt{Set}(x)$ *to be* $(\forall z)(\forall y \in x)(z =_{NFU} y \to \neg\neg z \in x))$;
$x \in_{NFU} y$ *to be* $Set(y) \wedge \neg\neg(x \in y)$.

$\texttt{Set}(x)$ is a condition on x rather like being stable but stronger. It might be helpful to reflect that my pet objects $\overline{B}x = \{y : x \notin y\}$ satisfy this condition, so it's not totally outlandish. It might also help to reflect that $\texttt{Set}(x)$ says precisely that x is a union of $=_{NFU}$-equivalence classes. Observe that if $\neg\texttt{Set}(x)$ then the target model believes x to be empty. This fits in with thoughts about extensionality.

These definitions are due to Holmes.

Observe even at this relatively early stage that

Lemma 95 $\texttt{Set}(x)$ *and* $x =_{NFU} y$ *together imply* $\texttt{Set}(y)$.

Proof:

Suppose $\texttt{Set}(x)$ (which is to say $(\forall z)(\forall w \in x)(z =_{NFU} w \to \neg\neg z \in x))$), and $\sim\sim x = \sim\sim y$.

We desire $\texttt{Set}(y)$, which is to say $(\forall z)(\forall w \in y)(z =_{NFU} w \to \neg\neg z \in y))$).

Let z be arbitrary, and $w \in y$ and $z =_{NFU} w$. Then $\neg\neg(w \in x)$ (since $\sim\sim x = \sim\sim y$).

Now $w \in x$ implies $\neg\neg(z \in x)$ so $\neg\neg(w \in x)$ also implies $\neg\neg(z \in x)$. So $\neg\neg(z \in y)$ too.

■

Naturally the intention is that the translation of all the axioms (indeed theorems) of NFU are theorems of *i*NF. (Indeed, one hopes that they will be theorems of *i*NFU!)

The precise details of the negative interpretation won't matter a great deal, and i'll not bother spelling them out for the moment. Comprehension is not going to be a problem. However, there are two things that are likely to be bottlenecks, and we must get them straight. Extensionality is one, and validating the rule of substitution is another.

14.1.1 Verifying the rule of Substitution

We want to be sure that if $x =_{NFU} y$ and ϕ then $[y/x]\phi$. That is, for ϕ in the language with $\texttt{Set}(-)$, $=_{NFU}$ and $\in_{NFU}$. (We don't need it for *all* ϕ!)

It will suffice to do this for atomics. [will it? I'm getting nervous .... Perhaps it will follow from substitutivity of the biconditional.]

There are three cases to consider: ϕ can be (i) $x =_{NFU} z$, (ii) $z \in_{NFU} x$, or (iii) $x \in_{NFU} z$.

> (i) is easy: $=_{NFU}$ is clearly transitive.
>
> For (ii) suppose $z \in_{NFU} x$ and $x =_{NFU} y$.
>
> We desire that $z \in_{NFU} y$, which is to say we desire $\mathrm{Set}(y) \wedge \neg\neg(z \in y)$
>
> $z \in_{NFU} x$ is $\mathrm{Set}(x) \wedge \neg\neg(z \in x)$
>
> and $x =_{NFU} y$ is $\sim\sim x = \mathit{doublecompy}$
>
> Now $\neg\neg(z \in x)$ and $\sim\sim x = \mathit{doublecompy}$ give $\neg\neg(z \in y)$ as desired. To get $\mathrm{Set}(y)$ we reflect that we have already shown (lemma 95) that $\texttt{Set}(x)$ and $x =_{NFU} y$ together imply $\texttt{Set}(y)$.
>
> For (iii) suppose $x \in_{NFU} z$ and $x =_{NFU} y$. Now $x \in_{NFU} z$ implies that $\texttt{Set}(z)$, so anything notnotequal to a member of z is also in z. But y is one such, so $y \in z$ whence $y \in_{NFU} z$.

Is it really sufficient to do it for atomics. . . ?

14.1.2 Verifying Extensionality

I am taking extensionality to be

$$(\forall x_1 x_2)(\forall y_1 y_2)(y_1 \in x_1 \wedge y_2 \in x_2. \to (\forall z)(z \in x_1 \longleftrightarrow z \in x_2) \to x_1 = x_2)$$

This says that two nonempty sets with the same extension are identical.

This means that we have to prove the following in *i*NF:

$$(\forall x_1 x_2)(\forall y_1 y_2)(y_1 \in_{iNF} x_1 \wedge y_2 \in_{iNF} x_2. \to (\forall z)(z \in_{iNF} x_1 \longleftrightarrow z \in_{iNF} x_2) \to x_1 =_{NF} x_2).$$

Or rather not literally that but the result of modifying it according to whatever rewritings of quantifiers and connectives is called for by our choice of clauses for the quantifiers and connectives.

Actually we want to show it in *i*NFU!

In the interpretation if x_1 and x_2 are two nonempty sets then we have $\mathtt{Set}(x_1)$ and $\mathtt{Set}(x_2)$ so they are both super stable sets and that ensures that if they have the same extensions they are identical.

14.1.3 Verifying Comprehension

The hard part is ensuring that $\mathtt{Set}(-)$ holds for any object that is a set abstrcat in this new sense.

The way to deal with that is to ensure that, for the ϕ that we are going to find extensions for, if $\phi(x)$ and $x =_{NFU} y$ then $\phi(y)$. That way the set of all things that are ϕ is a union of $=_{NFU}$-equivalence classes. So we need to prove substitutivity of equality for $=_{NFU}$—for negative expressions.

Again the details depend on what recursions we have adopted for the quantifiers and connectives. However the broad outlines of an explanation of why it works can easily be had. For any stratified expression ϕ in the language of set theory we consider its translation ϕ^* decorated with NFU subscripts and we want .... It has an extension (call it 'X') because ϕ^* is stratified as long as ϕ is. We want this X to be the extension of ϕ in the sense of *i*NF(U). Care is needed because $\in_{iNF}$ (at least when a nonempty set is to the right) $\neg\neg \in$ rather than $\in$. But any negative interpretation is going to have the effect that X is a stable set, and this ensures that $\in$ and $\neg\neg \in$ to the left of X are equivalent.

15 Holmes on Realizability for *i*NF

(I've doctored some of his notation—tf)

The constructive interpretation of what a proof is suggests an argument for the consistency of intuitionistic New Foundations.

Everything we deal with will be a term (a syntactical object of some sort). Some of these terms will be associated with functions of various kinds, but we will not be working with unrestricted function spaces.

We present a mutually recursive definition of what *propositions*, *proofs*, and *terms* are in intuitionistic T$\mathbb{Z}$T (simple theory of types with all integer types).

The False: $\perp$ is a proposition, which we hope has no proofs.

Conjunctions: If p and q are propositions and P and Q are proofs of p and q respectively, then $p \wedge q$ is a proposition and (P, Q) is a proof of this proposition.

Disjunctions: If p and q are propositions and P and Q are proofs of p and q respectively, then $p \vee q$ is a proposition and objects of the form $(\texttt{left}, P)$ and $(\texttt{right}, Q)$ are proofs of this proposition.

Implications: If p and q are propositions then $p \rightarrow q$ is a proposition and any function which maps each proof of p to a proof of q is a proof of this proposition. Lambda-abstraction over proofs with respect to proof variables is one way to construct such functions, and there will also be some special atomic functions to be described elsewhere in this definition. In any event, any proof which is a function is associated with some syntactical object.

Universal Statements: If p is a proposition and x is a variable of type i, $(\forall x.p)$ is a proposition and a function from terms a of type i to proofs of $p[a/x]$ (defined syntactically) is a proof of this proposition.

Existential Statements: If p is a proposition and x is a variable of type i, $(\exists x.p)$ is a proposition, and any pair (a, P), where a is a term of type i and P is a proof of $p[a/x]$, is a proof of this proposition.

Membership Statements: If x is a term of type i and p is a proposition, $\{x \mid p\}$ is a term of type i. There are terms of all integer types, and all terms other than variables of each type (of which we have as many as we need) are constructed in this way. If a and b are terms of type i and $i+1$, respectively, then $a \in b$ is a proposition. A proof of $a \in \{x \mid p\}$ is a proof of $p[x/a]$.

Equations: If a is a term of type i and b is a term of type i, $a = b$ is a proposition. This is actually definable as $(\forall x.a \in x \leftrightarrow b \in x)$. It needs to be noted that there is a special atomic proof `ext` which sends proofs of $(\forall x.x \in a \leftrightarrow x \in b)$ to proofs of $a = b$. Note that the biconditional can be defined in terms of conjunction and implication in the usual way.

All propositions, proofs and terms are defined syntactically. This is an implementation of intuitionistic or constructive T$\mathbb{Z}$T (though the notion of T$\mathbb{Z}$T as a constructively acceptable theory rather boggles the mind!)

Here, if anywhere, one should be able to exploit the original insight of Russell and Quine that the proof process is "typically ambiguous"; every proof corresponds to an equally valid proof in which all type indices are raised or lowered by a constant amount. If p is a proposition, call the type-raised version (with step 1) p^+. The map we have just referred to is a perfectly definite syntactical operation (which can be extended to terms and proofs as well). It would seem that this operation, which we might call `amb+` paired with its inverse, `amb-`, forms a proof of all propositions $p \leftrightarrow p+$ (instances of ambiguity). It appears that other type-shuffling functions will be needed to handle more general variants ($p \leftrightarrow p^+$ is not an adequate scheme of ambiguity for constructive logic), but the general idea holds: in all such cases, the existence of a proof with one set of types will enable us to construct a proof with the other set of types. This fact would not be perturbed by the addition of the new operators. Dzierzgowski points out in his thesis that this is a feature of the underlying first-order logic, not really a feature of type theory *per se*.

The questionable uses of ambiguity (the ones which deviate from results in classical logic) are those in which we apply ambiguity to a hypothesis. This seems valid: here is the reasoning. Suppose that we have introduced and not discharged a hypothesis p. This means that all our reasoning is under the assumption that we are given a proof of p. But we have stated above exactly what we mean by a proof of p, and it is clear that if we are given a proof of p we are also given a proof of p^+ (and *vice versa*)! The form of a "proof" would seem to be as follows: Con(TT) tells us that we don't get contradictions in the constructive theory; in the constructive theory, we restrict our reasoning under hypotheses to situations where we can be given a proof of the hypothesis; under these situations we also have proofs of variants of the hypothesis, so we are safe in using these proofs as well.

This argument is not valid, or at least more needs to be shown. Suppose that we did have a theorem which denied an instance of ambiguity (this actually won't happen, by a result of Crabbé which shows that any single instance is consistent with TT + AC, but it is a better example of my objection to (my own!) argument above). If we can prove $p \longleftrightarrow (p^+ \rightarrow \perp)$, this tells us that if we assume that we are given a constructive proof of p, we can construct a constructive proof of the negation of p with all type indices raised. This tells us, further, that there can be no constructive proof of p or of its negation (by the same considerations of ambiguity stated above) unless TT is inconsistent. But it is not at all clear that a proof of this kind would enable us to prove in T$\mathbb{Z}$T that p could not have a constructive proof (that is, prove that from a proof of p we can construct a proof of $\perp$). It *would* seem to imply that T$\mathbb{Z}$T, and, indeed, TT itself, were quite perverse from a constructive standpoint.

This line of thought would complete itself to a proof of Con(*i*NF) if we could establish some kind of result to the effect that any statement which can't be disproved in TT "can" have a constructive proof in TT in some sense. It would complete itself to a contradiction in Con(*i*NF) (and so in NF itself) if we could find some p (a kind of Gödel sentence) which actually had the behaviour described above (it would have to be a little more complex)! This last outcome is not impossible to imagine; we have already seen sentences which assert their own unprovability. The idea above is to weaken "true" to "has been proven". This is not good in some cases: a Gödel sentence may assert that it cannot be proved.

Investigate the construction of notions of constructive proof as above inside T$\mathbb{Z}$T or *i*NF. Are there reflection results which could be used one way or the other?

15.1 tf on realizability for Constructive T$\mathbb{Z}$T

We have no free variables. Our atomics are $t_i = t_j$ and $t_i \in t_j$. What is a realizer for $t_i = t_j$? It's a pair of a realizer for $t_i \subseteq t_j$ and a realizer for $t_j \subseteq t_i$. And a realizer for $t_i \subseteq t_j$ is of course a function that takes a term t_k and returns a function that takes realizers for $t_k \in t_i$ and returns realizers for $t_k \in t_j$.

A realizer for $t_i \in t_j$ is of course a realizer for $\phi(t_i)$, where t_j is $\{x : \phi(x)\}$.

Manifestly this process is *prima facie* illfounded! It should be fairly straightforward to find explicit nonterminating descending chains.

I have the feeling that i must have misunderstood this very badly. What do we have in the way of realizers for $t_i = t_j$ if the two terms happen to e syntactically identical? We certainly want the identity function. Does our recursion give us that?

16 Correspondence with Daniel

From Daniel

As I told you, I'm studying what I call N-finite sets, where the empty set is N-finite and if x is N-finite and y is not in x, then $x \cup \{y\}$ is N-finite. This allows [us] to interpret arithmetic (up to now in ITT with a suitable axiom of infinity; I think I can do it in ITT_3 as well). But we must be careful. For example, I discovered yesterday than my favorite axiom of infinity

$(\forall x)$ N-finite $\exists y y \notin x$

is strictly stronger than

$\forall x, y$ N-finite $\exists x', y'$ N-finite s.t. $x \sim x'$ and $y \sim y'$ and $x' \cap y' = \emptyset$.

($\sim$ means "there's a bijection from ... to ...). Funny, isn't it ? Also, $x \sim y$ is not equivalent to $(x \setminus y) \sim (y \setminus x)$, but is equivalent to $\neg\neg(x \setminus y) \sim (y \setminus x)$.

*i*NF is equiconsistent with ITT + $\phi \longleftrightarrow \phi^*$, isn't it?

First remark that the excluded middle for (possibly open) weakly stratified formulæ is a consequence of the excluded middle for stratified formulae (if you have the E.M. for all open stratified formulae, than you have the universal closure of those E.M., and E.M. for weakly stratified formulae appears when you eliminate the universal quantifiers). Then *i*NF + EM for stratified formulae is equiconsistent with ITT + ambiguity + EM for all formulae, and thus equiconsistent with classical NF. This seems to be correct, isn't it ?

From Daniel

From ddz@agel.ucl.ac.be Wed Jun 8 13:54:13 1994

AxInf is the axiom of infinity given in my draught about finite sets, i.e. $(\forall x \in NFin)(\exists y \in NFin)(y > x)$ ($>$ means there's a 1-1 function mapping x into y and no 1-1 function mapping y into x).

This AxInf is also equivalent to $(\forall x \in NFin)(\exists y \in NFin)(x \sim USC(y))$

Both NFin are not of the same type, of course.

I've included a copy of my lost message below. The counter-model there mentionned is in fact simpler than I first thought. I'll explain later.

Daniel

From Daniel

I've just looked at the relation between AxInf and AxInf$^+$ in ITT ($^+$ means "raise types" as usual). In classical TT, AxInf and AxInf$^+$ are equivalent (is this important, by the way ?).

In ITT, it's quite easy to prove that AxInf implies AxInf$^+$. Anyway, the converse doesn't seem to hold. I've got an idea for a counter-example. I have not yet written it down, but I think it's correct. The idea is to take a Kripke model $\mathfrak{M}$ whose domains of type 0 are all equal to $\{x_0, x_1, x_2, ...\}$ and such that $\mathfrak{M}$ does *not* satisfy $(x_i = x_j \vee x_i \neq x_j)$ if $i \neq j$ (I can find such an $\mathfrak{M}$). Then, in $NFin^1$, there are only the empty set and singletons. There are no pairs $\{x_i, x_j\}$ in NFin^1, because x_i should be $\neq x_j$, which is never the case. So AxInf is not satisfied for NFin^1. Anyway, I think AxInf is satisfied for NFin^2. NFin^2 contains N-finite sets which are as great as you want. I think I can find N-finite sets whose members are in SFin^1, and which are great. I'll write down the details; that could fill one page or two. If you can find a simpler counter-example,... or a proof of $AxInf \longleftrightarrow AxInf^+$... I'd be glad to read it.

From Daniel

Hello !

How are you doing with intuitionistic TT and NF? Anything new?

Here are some ideas. I guess you should know about all this.

1. The initial idea of Marcel was to try to build a Kripke model of *i*NF. He wanted to start from a classical model of NFU. Then, by means of permutation methods, you can remove urelemente. This gives you a "better approximation" of a model of NF. But we cannot remove all the urelemente (I mean, we don't know how to do...). So Marcel's idea was to arrange all these better and better approximations of a model of NF in a Kripke model of intuitionistic NF. This sounds great. Unfortunately, it doesn't work at all. Try it out and you will quickly understand.

2. In a model of (a fragment of) TT or NF, there is at least one singleton whose only subsets are the empty set and the singleton, then the excluded middle is satisfied in the model.

3. I think you should work in the theory before trying to prove it is consistent. For example, you could study arithmetic or the axiom of infinity (see below).

4. I believe toposes are worth studying. I have put a short file (in French) explaining roughly how to interpret intuitionistic TT within a topos at http://users.skynet.be/ddz/nf.html. I've discovered that the technique I explained in my JSL paper can be rephrased more clearly in terms of toposes (I mean, it is clearer if toposes are clear for you...). People working on toposes know many example of toposes. Maybe there is one we would no thought about and that could help to build some useful model of TT.

5. Proof theory is certainly a interesting point of view to study the consistency of int. NF. I didn't investigate it at all.

6. I'm not sure that weak ambiguity is of some interest. It does not have nice property from a proof theoretical point of view. I don't know what you can prove using weak ambiguity.

7. I have studied arithmetic with much care. You can get a draft paper with some more properties of arithmetic in int. TT from http://users.skynet.be/ddz/nf.html. The file I've mentioned above, about toposes, gives an hint to prove that the arithmetic of int. NF is not classical (if int. NF is consistent). I beleive it is a nice problem to study; it is quite complexe, but one should be able to work it out.

8. In my study of arithmetic, I've proved that some classically equivalent forms of the Axiom of Choice are no more equivalent in an intuitionistic framework. I've pointed out the form that is adequate for the interpretation of arithmetic. But I don't know how good it is for other purposes. I don't either know how to prove it in int. NF. This is also a nice problem to study.

   I can explain more. Just ask!

Best wishes, Daniel.

17 Wellfounded sets

The inductive definition of wellfounded sets is the obvious one: if we think of well-founded sets as those over which we can do $\in$-induction then we are led to the inductive definition:

$$WF(x) \longleftrightarrow (\forall y)((\forall z)(z \subseteq y \to z \in y) \to x \in y) \tag{9}$$

The class of wellfounded sets is the least fixpoint for the power set operation. This definition is legitimate in both the classical and constructive versions of NF. It's not legitimate in ZF because ZF proves that no set extends its own power set. There is an "upside-down" definition of wellfounded set which is available in ZF, but the equivalence of the two definitions needs excluded middle, and is accordingly not available in *i*NF ... but then it isn't needed! *i*NF has a smooth treatment of wellfounded set. Observe that the definition in 9 is not stratified.

And it cannot be made to be stratified, since o/w we would have Mirimanoff's paradox.

With this definition we can justify $\in$-induction for wellfounded sets for stratifiable expressions.

18 Nearly solving the dilemma

The dilemma: does *i*NF interpret Heyting Arithmetic?

Suppose $V/(\neg\neg =)$ is kfinite. (This is on the face of it a slightly stronger assumption than the existence of a dense Nfinite set). Then it has a transversal. Call this transversal T. The existence of such a T implies weak de Morgan (see 72). We know $(\forall x)(\exists y \in T)(\neg\neg(x = y))$. (This follows specifically beco's T is a transversal: merely be dense Nfinite is not enuff). Suppose $\neg(\exists x \in T)F(x)$; then $(\forall x \in T)\neg F(x)$. Then $(\forall x)\neg F(x)$. Contraposing, if we assume $\neg(\forall x)\neg F(x)$ we infer $\neg\neg(\exists x \in T)F(x)$. But now (since T is kfinite) we can use the $\exists$ version of Linton-Johnstone that follows from weak de Morgan (that was lemma 46) to push the $\neg\neg$ inside and infer $(\exists x \in T)\neg\neg F(x)$, which gives $(\exists x)\neg\neg F(x)$. So we have proved

$$\neg(\forall x)\neg F(x) \to (\exists x)\neg\neg F(x).$$

This is a nontrivial quantificational principle!

But it also means you can import $\neg\neg$ past $\exists$! If $(\exists x)F(x)$ then $(\exists x \in T)\neg\neg F(x)$. Wrap it up in $\neg\neg$ to get If $\neg\neg(\exists x)F(x)$ then $\neg\neg(\exists x \in T)\neg\neg F(x)$. But now, since T is finite, we can import the $\neg\neg$ by Linton-Johnstone, getting $(\exists x \in T)\neg\neg F(x)$, which of course implies $(\exists x)\neg\neg F(x)$.

But this actually follows from $\neg\forall\neg \to \exists\neg\neg$ beco's $\neg\neg\exists \to \neg\forall\neg$ ... beco's $\forall\neg \to \neg\exists$

[I was hoping it may even mean that we can *export* $\neg\neg$ past $\forall$ ... Suppose we have $(\forall x)\neg\neg F(x)$. Then we have $(\forall x \in T)\neg\neg F(x)$. Linton-Johnstone now gives $\neg\neg(\forall x \in T)F(x)$. Now suppose *per impossibile* we have $\neg(\forall x)F(x)$ err...]

So, if $V/(\neg\neg =)$ is kfinite, we infer THREEE logical principles:

(1): $\neg\neg(A \vee B) \ = \ \neg\neg A \vee \neg\neg B$;

(2): $\neg(\forall x)\neg F(x) \to (\exists x)\neg\neg F(x)$;

(3): $\neg\neg(\exists x)F(x) \to (\exists x)\neg\neg F(x)$.

2 implies 3. I haven't managed to prove that $\neg\neg$ commutes with $\forall$, tho' i don't see why that case should be any different from $\exists$.

To get some action we need to show that $\neg(\forall x)\neg F(x) \to (\exists x)\neg\neg F(x)$ means that $V/(\neg\neg =)$ cannot be finite.

It would be nice, too, if we could show that the existence of a dense Nfinite set implies that $V/(\neg\neg =)$ is finite.

One would half-expect, too, that principle (2) above should imply that the logic is classical.

But hang on! We proved in corollary 18 that $V/(\neg\neg =)$ was not finite!!

18.0.1 I tho'rt I'd written this down but perhaps I hadn't

Sse $\mathcal{V}$ is a dense Nfinite set. (Actually i think we need $V/(\neg\neg =)$ to be kfinite and $\mathcal{V}$ to be a transversal for it). Then we get Peter's weak de morgan and the analogue of Linton-Johnstone for $\exists$.

Suppose $\neg\neg(\exists x)F(x)$

If $F(x)$ then let a be something s.t $F(a)$. Then $\neg\neg(a \in \mathcal{U} \wedge F(a))$, whence $(\exists x)\neg\neg(x \in \mathcal{U} \wedge F(x))$. We can export the $\neg\neg$ to get $\neg\neg(\exists x)(x \in \mathcal{U} \wedge F(x))$.

So $(\exists x)F(x)$ implies $\neg\neg(\exists x)(x \in \mathcal{U} \wedge F(x))$, giving $\neg\neg(\exists x)F(x)$ implies $\neg\neg\neg\neg(\exists x)(x \in \mathcal{U} \wedge F(x))$ which is $\neg\neg(\exists x)(x \in \mathcal{U} \wedge F(x))$.

Then, by Linton-Johnstone for $\exists$, we infer $(\exists x)\neg\neg(x \in \mathcal{U} \wedge F(x))$ and finally $(\exists x)\neg\neg F(x)$.

Thus $\neg\neg(\exists x)F(x) \to (\exists x)\neg\neg F(x)$.

But that STILL doesn't seem to be enough

18.0.2 Dense Nfinite sets again

Can we exploit the fact that $V/(\neg\neg =)$ is not kfinite to prove that there is no set $\mathcal{V}$ that is dense and Nfinite? We have seen (remark 43) that if there is such a $\mathcal{V}$ then it is unique to notnot-equality. The lacuna is the fact that

$$(\forall x)\neg\neg(x \in \mathcal{V})$$

does not imply

$$(\forall x)(\exists y \in \mathcal{V})(\neg\neg(x = y))$$

If that inference were good then the map $x \mapsto [x]_{\neg\neg=}$ would be onto and we could conclude that $\mathcal{V}$ is not kfinite.

But consider $\{[y]_{\neg\neg=} : y \in \mathcal{V}\}$. This set is kfinite and so cannot be equal to $V/(\neg\neg =)$ (which isn't); but it's unique up to $\neg\neg =$; it doesn't depend strongly on $\mathcal{V}$.

$V/(\neg\neg =) \setminus \{[y]_{\neg\neg=} : y \in \mathcal{V}\}$ is nonempty, but of course is not inhabited.

So what about $\{y : (\exists z \in \mathcal{V})\neg\neg(y = z)\}$? Can we put it to use? Again, it is unique to notnot-equality. And it's 1-symmetric, tho' that probably doesn't help.

18.1 Intuitionistic wellorderings as sets of initial segments

It is a little-known but standard and unproblematic fact that wellorderings can be stored as the set of (carrier sets of) their terminal segments. Let us say

Definition 96 *$\mathcal{X}$ is a* **wellordering** *of X iff*

1. $(\forall x, y \in \mathcal{X})(x \subseteq y \vee y \subseteq x)$
2. $(\forall \mathcal{X}')(\mathcal{X}' \subseteq \mathcal{X} \to \bigcup \mathcal{X}' \in \mathcal{X}')$

3. $(\forall x_1, x_2 \in X)((\forall X' \in \mathcal{X})(x_1 \in X' \longleftrightarrow x_2 \in X') \to x_1 = x_2)$

4. $\bigcup \mathcal{X} = X$

Allen Hazen calls these things **ordernestings**.
If X has such an $\mathcal{X}$ we say X is **wellordered**.

Check this definition: i'm not sure that i've got it right. I know *something* like this works ...

It is easy to show that any two wellorderings are comparable in point of length. Given two wellorderings $\mathcal{X}$ and $\mathcal{Y}$ we have the inductively defined set which pairs the empty set with the empty set, and whenever it contains a bijection between $\mathcal{X}' \subseteq \mathcal{X}$ and $\mathcal{Y}' \subseteq \mathcal{Y}$ it also pairs $\bigcap\{A \subseteq \mathcal{X} : (\forall X' \in \mathcal{X}')(\neg(A \subseteq X'))\}$ with $\bigcap\{B \subseteq \mathcal{Y} : (\forall Y' \in \mathcal{Y}')(\neg(B \subseteq Y'))\}$.

Every subset of a wellordered set is wellordered. If $\mathcal{X}$ is a wellordering of X, and $X' \subseteq X$, then $\{A \cap X' : A \in \mathcal{X}\}$ is a wellordering of X'.

Every quotient of a wellordered set is wellordered. If $\mathcal{X}$ is a wellordering of X, and f is defined on X then $\{f``X' : X' \in \mathcal{X}\}$ is a wellordering of $f``X$.

We prove by induction that every Kfinite set has a wellordering: If $\mathcal{X}$ is a wellordering of X, then $\mathcal{X} \cup \{X \cup \{y\}\}$ is a wellordering of $X \cup \{y\}$. The converse is not true. Indeed one can have wellordered subfinite sets that are not Kfinite, for example $\{x : x = a \wedge p\}$ is wellordered, because its singleton is a wellordering of it, but it is not Kfinite unless $p \vee \neg p$.

It does not follow from $\mathcal{X} \sim \mathcal{Y}$ that $|X| = |Y|$! $\mathcal{X}$ and $\mathcal{Y}$ might contain silly subsets of X and Y.

If $\mathcal{X}$ is a wellordering of X then $\lambda x. \bigcap\{Y \in \mathcal{X} : x \in Y\} : X \hookrightarrow \mathcal{X}$. The obvious inverse sends $X \in \mathcal{X}$ to the unique member of X-minus-the-union-of-all-its-proper-subsets-in $\mathcal{X}$, but there is no reason to expect this to be defined.

Notice that $\mathbb{N}$ is wellordered. We can define $\leq_{\mathbb{N}}$ as an inductively defined set. Then the collection of initial segments is a set.

Similarly there is no cut-free proof of the sequent $(\exists x)(\forall y)(y \in x) \vdash$.

Ordernestings of Nfinite Sets

We can, in a stratified manner, define a datatype of ordernestings of Nfinite sets: if O is an ordernesting of X, and $x \notin X$, then $O \cup \{X \cup \{x\})$ is an ordernesting of $X \cup \{x\}$.

This declaration enables us to prove by Nfinite induction that every Nfinite set has an ordernesting.

Can we show that the set of ordernestings of an Nfinite set is Nfinite? Nothing to say we can't. How would the proof go? There is work to do beco's $X \cup \{x\}$ is Nfinite in lots of different ways. It would help if we could show that Symm(X) is kfinite (Nfinite) as long as X is.

Can we then go on to prove that if $\mathcal{V}$ is an Nfinite dense set then every $\overline{B}x$ is a double complement of an Nfinite subset of $\mathcal{V}$? Then we would be close to a proof that $\overline{B}``V$ is kfinite and then things really start to hum.

Can we show that any two ordernestings an a given Nfinite set are iso? Every ordernesting of an Nfinite set is Nfinite?

19 Stuff to be tidied up

There is a question about whether or not the set of cardinals of Nfinite sets is finite.

Presumably one can prove that if x is an Nfinite set then ι^2“x is in 1-1 correspondence with the Nfinite cardinals $< |x|$?

If so, then, for any Nfinite cardinal n of the form $T^2 k$ for some k, equality between numbers below n is decidable.

Albert points out that there is a question about what the propositional logic of *i*NF is. It's not at all obvious that it has to be ordinary constructive propositional logic. After all, set theoretic principles can imply logical principles—think of AC implying excluded middle!

Suppose $\mathcal{V}$ is dense and Nfinite. Suppose $a \neq b$ both in $\mathcal{V}$. Let f be a fishy combination of a and b, so that $\neg(f \neq a \wedge f \neq b)$. If $f \in \mathcal{V}$ then we have both $f = a \vee f \neq a$ and $f = b \vee f \neq b$. Distributing we get four disjuncts

$$f = a \wedge f = b,\ f \neq a \wedge f \neq b,\ f = a \wedge f \neq b \text{ and } f \neq a \wedge f = b.$$

The first two are impossible so we conclude $(f = a \wedge f \neq b) \vee (f \neq a \wedge f = b)$. So the only fishy combinations of a and b that are in $\mathcal{V}$ are a and b. But every fishy combination of a and b is notnot in $\mathcal{V}$. So every fishy combination of a and b is notnot equal to either a or b. But we knew that anyway. Bah.

We probably need to show in *i*NF (and it might be a good idea to do this in NF too) that the set Nfin/equipollence (i) is the same as the inductively defined set (ii) containing $\{\emptyset\}$ and closed under $X \mapsto \{y \cup \{z\} : z \notin y \in X\}$.

Both directions would be inductions. We would prove by induction on (ii) that all its members are equipollence classes of Nfinite sets, and we would prove by Nfinite induction that the equipollence class of an Nfinite set belongs to (i).

Now there is an interesting complication, in that (i) is a partition and cannot contain the empty set, whereas (ii) in principle might contain an empty set.

WK pairs are constructively robust

Somewhat to my surprise, the (Wiener)-Kuratowski pair turns out to be constructively robust. This is quite striking beco's usually one does a case split on whether or not the two components are identical. Thanks to PTJ and Randall Holmes who both spotted it. It seems to me that this *aperçu* is probably worth writing out in some detail.

Let p be a pair, which is to say $(\exists x, y)(p = \{\{x\}, \{x, y\}\})$. Then $\mathtt{fst}(p)$ is $\bigcap\bigcap p$ (that's easy) and $y = \mathtt{snd}(p)$ iff $y \in \bigcup p$ and $(\forall u, v \in p)(y \in u \wedge y \in v. \rightarrow u = v)$.

Let us now assemble a Kuratowski pair p' whose two components are the two components we have extracted from p. We have to establish that $p = p'$.

Naturally our weapon is extensionality: we must show that p and p' have the same members. I think we can agree that $\{x\}$ belongs to both p and p'.

$p \subseteq p'$;

Things in p are either singletons or doubletons. If you are a singleton you are either $\{x\}$ (so you are in p) or you are $\{x, y\}$ (in which case $x = y$) so once again you are $\{x\}$ and you are in p.

If you are a doubleton you are $\{x, y\}$ with $x \neq y$, and then again you are in p.

The other direction is similar.

But we do need more detail!

Being a set theoretic foundationalist is of course crazy, and it doesn't become any more crazy (or any less crazy) if you adopt a constructive stance. But adopting a constructive stance concentrates the mind on features that one might miss if one's take is purely classical.

Definitions/constructions of inductively defined objects. Top-down vs bottom-up. Top-down v easy if you have big sets (or even intermediate sets)

We know that commutation of $\forall$ with $\neg\neg$ is as strong as NF, so it certainly implies has at least the consistency strength of an ability to interpret Heyting arithmetic. It would be nice to have a direct proof. We want to prove

$$\neg\neg(\forall x)(Nfin(x) \to (\exists y)(y \notin x))$$

which is what we need if we are to interpret Heyting arithmetic.

We can get it from

$$(\forall x)\neg\neg(Nfin(x) \to (\exists y)(y \notin x)).$$

Now $A \to \neg\neg B$ implies $\neg\neg(A \to B)$ so we want

$$(\forall x)(Nfin(x) \to \neg\neg(\exists y)(y \notin x)),$$

because that will imply

$$(\forall x)\neg\neg(Nfin(x) \to (\exists y)(y \notin x))$$

and we can pull the $\neg\neg$ out (even without commutation) to obtain

$$\neg\neg(\forall x)(Nfin(x) \to (\exists y)(y \notin x)).$$

So let's try and prove

$$(\forall x)(Nfin(x) \to \neg\neg(\exists y)(y \notin x))$$

Now! If $Nfin(x)$ we have $x \neq V$, from which we infer $\sim\sim x \neq V$ by commutation. But this is $\neg(\forall y)\neg\neg(y \in x)$, and we know that $\neg\forall\neg\neg$ implies $\neg\neg\exists\neg$

Crabbé proved the consistency of two predicative fragments of NF: NFP and NFI. Both are extensionality plus existence of $\{y : \phi(y, x_1 \ldots x_n)\}$ where ϕ is weakly stratified and there are extra restrictions on the variables in ϕ. All variables (bound or free) are restricted to type no higher than that of the set being defined; parameters are not to be of any type exceeding the type of y (NFP) or (type of y) + 1 (NFI). NFP + Axiom of sumset = NF. In NFP we argue: either the axiom of sumset holds, in which case we have NF and can run Specker's proof of the Axiom of infinity, or (ii) it doesn't, in which case—since finite sets have sumsets—not every set is finite.

How does this play constructively? Well, it's more-or-less the same. Suppose every set is subfinite. In *i*NFP we prove by induction on the kfinite sets that every subset of a kfinite set has a sumset. To do this we have to show that the set $\{x : \bigcup x \text{ exists}\}$ exists and is kfinitely closed so-to-speak. Fortunately the existence of this set is an axiom of *i*NFP. So every subfinite set has a sumset. But if V is kfinite then every set is subfinite and has a sumset and we are in *i*NF and can prove that V is not kfinite.

■

Does that mean that if *i*NF interprets Heyting Arithmetic then so does *i*NFP?

We can interpret classical TZT into the constructive version by adding to $L(\mathrm{T}\mathbb{Z}\mathrm{T})$ a constant symbol $\mathfrak{P}_n$ at each level n and restricting all our variables to those constants. I think that works. A more subtle and complicated question is: can we modify this to interpret T$\mathbb{Z}$T+Amb into iT$\mathbb{Z}$T+Amb? This amounts to requiring that we prove $\phi^{\mathfrak{P}_n} \longleftrightarrow \phi^{\mathfrak{P}_{n+1}}$ in iT$\mathbb{Z}$T+ Amb. This is not straightforward because there are constant symbols in $\phi^{\mathfrak{P}_n} \longleftrightarrow \phi^{\mathfrak{P}_{n+1}}$ and the ambiguity scheme does not cover formulæ containing constants. We can expect iT$\mathbb{Z}$T to prove $\phi^X \longleftrightarrow \phi^{\mathfrak{P}(X)}$ (where $\mathfrak{P}()$ is the Powell power set) unless there is something special about X. I see no hope of proving anything like that,

Of course the possibility of inner models for CO theories means there can be negative interpretations for CO theories. If there is no negative interpretation for NF that lends support to Kaye's conjecture.

Is there reason to hope that if *i*NF interprets Heyting arithmetic then so does *i*NFP? None evident to me. It may be that *i*NF refutes the possibility of a dense Nfinite set but only by using some impredicative set abstracts.

Is the equivalence of stratified formulæ with acyclic formulæ good constructively?

Here's something that should be spelled out clearly for beginners ... We know that the axiom of choice is provable for finite families, but we also know

Probably not in the right place

that the axiom of choice for pairs implies excluded middle. Isn't this a contradiction? No. The point is that the (classical) inductive proof of AC for finite families goes over into a constructive proof that Nfinite families have choice functions; the pairs that feature in Diaconescu's proof (and there seem to be several versions around, all involving pairs) are Kfinite but not obviously Nfinite unless excluded middle holds.

19.1 Constructive TTT

I used to think that the fact that there is no constructive demonstration that there is no surjection $X \twoheadrightarrow (X \to X)$ opens the door to a construction of a constructive model of Tangled Type Theory. However life is a bit more complicated beco's there can be such a surjection only if X does not have two distinct elements. So there will be no surjection $X \twoheadrightarrow (X \to \Omega)$.

Actually we can show that there is no surjection $X \twoheadrightarrow (X \to X)$ as long as there is $f : X \to X$ with no fixed point. Suppose $g : X \to (X \to X)$. Then $\lambda x.f((gx)x)$ is not in the range of g. For suppose it is $g(a)$. Then

$$(g(a))(a) \;=\; \lambda x.f((gx)x)a \;=\; f((ga)a) \neq g(a)(a).$$

But surely "X has two distinct elements" follows from "$\exists f : X \to X$ wth no fixed point"? If X is inhabited, with $x \in X$, then $f(x)$ is another inhabitant, and $x \neq f(x)$.

Then reflect that if X is inhabited (by a, say) and $\neg(\forall x, y \in X)(x = y)$ then $\mathbf{1}_X$ and $\lambda x.a$ are distinct.

$X \to X$ is always inhabited—by $\mathbf{1}_X$ even if by nothing else. Now suppose it does not have two distinct elements. If X is inhabited then we obtain a contradiction.

So constructively we can prove that if X maps onto $X \to X$ then X has precisely one element.

Suppose $g : X \twoheadrightarrow (X \to X)$ and that a, b are two members of X. We will show that $a = b$... he says optimistically. $X \to X$ contains $\mathbb{1}_X$, $\lambda x.a$ and $\lambda x.b$

But perhaps all we can prove is $\neg\neg(a = b)$.

Perhaps the thing to do is to prove that there is no surjection $X \twoheadrightarrow ((X \to X) \to (X \to X))$.

Then the project of a constructive model of TTT really *will* be doomed.

Another take on the same material

Carsten Butz has an idea that the weakness of the versions of Cantor's theorem available constructively might make it possible to get constructive versions of Holmes' tangled type theory from [25] and thereby a model of *i*NF. This is probably worth explaining in some detail.

The point of departure here is the Boffa-Jensen proof of Con(NFU) by means of extracted models. When $\mathcal{M}$ is a structure for the language of simple type theory we **extract** a model $\mathcal{M}'$ from it by discarding some of the levels of $\mathcal{M}$ and defining new membership relations between the levels that are left. For example if we discard levels 2 and 3 we have to find a way of defining a membership relation between levels 1 and 4. This causes a problem because typically there will be more things in level 4 than can be distinguished (extensionally) by their "members" at level 1. In the original treatment one coped with this by just ruling that most things of (the old) level 4 are to be *urelemente* in the new dispensation. That is why this technique gives models of NFU not NF. Of course if $\mathcal{M}$ is a natural model (one where level $n+1$ is genuinely the power set of level n) then one has to have recourse to some such ruse. However there are plenty of models in which all the levels are the same size (seen from outside) so there is in principle the possibility of finding a new *extensional* membership relation between levels 1 and 4 if $\mathcal{M}$ is such a model. However, even in non-natural models $\mathcal{M}$ one doesn't expect there to be such a membership relation that is definable in $\mathcal{M}$ but there is no obvious impediment to simply expanding $\mathcal{M}$ by adding such a relation.

How strong a condition is it on $\mathcal{M}$ that one should always be able to find new membership relations for extracted models in this way? Holmes [25] has a very nice result that says that **if** $\mathcal{M}$ is such that whenever we extract a model $\mathcal{M}'$ from $\mathcal{M}$ by discarding some of the levels of $\mathcal{M}$ there is always a way of defining a membership relation between successive retained types that is nevertheless extensional **then** $\mathcal{M}$ gives rise to a model of NF.

Holmes uses the word **tangled** to describe models that allow extractions in this way. In the classical case it is *prima facie* unlikely that one could cook up these extensional relations between distantly separated types, because the fact that there are only two truth-values constrains the sizes of the various types quite closely. Indeed in the case of natural models an easy cardinality argument shows it to be outright impossible. If we have two sets A and B with $|B| \geq 2^{2^{|A|}}$, and $R \subseteq A \times B$ we will find that there are too many things in B for us to be able to distinguish b and $b' \in B$ by comparing $R^{-1}``\{b\}$ and $R^{-1}``\{b'\}$. So we shouldn't be surprised that the assumption that it is nevertheless possible (albeit for non-natural models) should turn out to be strong.

That makes the assumption that one can perform such cookery a strong one in the classical case, so it isn't surprising that one can find a model of NF by starting with a tangled model of Type theory.

However in the constructive case one has more *slop*. (This wonderful expression is Holmes'.) The problem was always that if there is to be an extensional relation between level 2 and level 4 then the size of level 4 must be 2-to-the-power-of the size of level 2 (or appear to be)—2 being the number of truth-values. Perhaps if we allow more truth-values we acquire more room for manœuvre? For example, rather than resign ourselves to consigning most elements of level 4 to being mere *urelemente* in the extracted model when we discard levels 2 and 3, we accommodate the fact that most of them will be hard to tell apart by conceding that an awful lot of x and y at type 4 are not properly distinct from one another?

However it now seems to me that constructively the situation is not as different from the classical case as one could have wished. The problem with the classical case is that if we extract levels 0 and 2, and discard level 1, we have to find an extensional relation between level 0 and level 2, and this forces level two to be the same size as level 1. Although this is not impossible, it prevents the model being a natural one. Unfortunately exactly the same argument works in the constructive case, so we are not on the face of it any further forward.

19.2 An old idea

I conclude with an idea for showing the consistency of *a kind of* intuitionistic NF. I'm sure other people have had this idea, but nobody seems to have discussed it in print, and it does not seem to be known whether it achieves anything or not.

Suppose $\langle V, \in, = \rangle$ is a model of NFU. We construct a Kripke model as follows. It will be an ω-sequence of worlds, and we characterise them as follows. Every W_i has domain V, and it has a membership relation $\in_i$ and equality $=_i$.

- W_0 is $\langle V, \in = \rangle$
- $x =_{i+1} y$ iff$_{df}$ $W_i \models (\forall z)(z \in x \longleftrightarrow z \in y)$ and $x \in_n y$ iff$_{df}$ $(\exists w)(x =_n w \wedge w \in y)$

It should be clear what this is an attempt to do. None of the W_i satisfy extensionality, but the sequence of models corresponds to an attempt to define a "contraction" by recursion, which of course won't work because V is not wellfounded.

It seems to me that the result is a Kripke model of something like an intuitionistic version of NF. In particular—as far as I can see—we have

$$(\forall xy)((\forall z)(z \in x \longleftrightarrow z \in y) \to \neg\neg(x = y))$$

Surely someone must have thought of this before? Where have I gone wrong?
Thomas

Daniel replies:
Here's the point. Consider A and B, two distinct empty sets in V, the model NF U. If the Kripke model satisfies the comprehension schema, it should satisfy the existence of a set E such that

$$(\forall z)(z \in E \longleftrightarrow A = B).$$

Thus, E is empty in W_0 and E is the universe in all W_{i+1}. I think that such an E does not exist in your construction.

Consider the two stratified assertions

$$(\forall y \in x)(\neg(\forall z)(z \in y)) \to \neg(\forall z)(z \in x) \qquad \text{A}$$
$$(\forall y \in x)((\exists z)(z \notin y)) \to (\exists z)(z \notin x) \qquad \text{B}$$

A has a constructive proof. In fact it has a constructive cut-free (normal) proof and a constructive stratified proof but—as far as i know—no stratfified cut-free (normal) proof. B has no constructive proof known to me.

19.2.1 A small joke

Recall the definition of an orthogonal set: a set x is **orthogonal** iff $(\forall yz \in x)(\neg\neg(y = z) \to y = z)$

REMARK 97 *Suppose iNF has the existence property. Then* $iNF \not\vdash (\forall x)(orthogonal(x) \to (\exists y)(y \notin x))$.

Proof: Suppose not, and that *i*NF both has the existence property and proves $(\forall x)(\text{orthogonal}(x) \to (\exists y)(y \notin x))$. Notice that if x is orthogonal and $y \notin x$ then $x \cup \{y\}$ is orthogonal too. Also a nested (indeed a *directed*) union of orthogonal sets is orthogonal. By assumption $iNF \vdash (\forall x)(\text{orthogonal}(x) \to (\exists y)(y \notin x))$. Therefore, for some term t, $iNF \vdash (\forall x)(\text{orthogonal}(x) \to (t_x \notin x)$. Notice that because of the equivalence with type theory, this t_x must be one type lower than x (otherwise this would not be a theorem of the underlying intuitionistic type theory!) But this enables us to build an unbounded increasing sequence of orthogonal sets and derive a Burali-Forti style paradox as follows. Let

> $B = \bigcup\bigcap\{Y : (\forall x \in Y)(x \cup \{t_x\} \in Y \wedge (Y \text{ is closed under nested unions}))\}$

B is clearly a paradoxical set, being a maximal orthogonal set. $B \cup \{t_B\}$ would also be orthogonal. This contradiction proves the remark. ■

If *i*NF $\vdash (\forall x)(\text{orthogonal}(x) \to (\exists y)(y \notin x))$ then certainly *i*NF proves that V is not orthogonal, so certainly classical NF will be inconsistent. Might remark 97 therefore mean that if *i*NF has the existence property then NF is consistent? Here we have to be very careful. Granted: if *i*NF has the existence property and NF is not consistent then *i*NF$\not\vdash (\forall x)(\text{orthogonal}(x) \to (\exists y)(y \notin x))$. So what happens if we invent a new constant a and claim orthogonal(a) and $\neg(\exists y)(y \notin a)$? Does this result in inconsistency? If it does, it means we have a proof in *i*NF that $(\forall x)(\text{orthogonal}(x) \to \neg\neg(\exists y)(y \notin x))$. But that might still be the case!

Indeed it is a lot more complicated even than that. If i understand Dzierzgowski right, then even adding the scheme $\phi \longleftrightarrow \phi^+$ to intuitionistic typed set theory is not sufficient to get a system equiconsistent with *i*NF. One needs a scheme $\phi \longleftrightarrow \phi^*$ where ϕ^* is obtained from ϕ by raising variables by any number of types one wishes: that is to say, one can raise different variables by different amounts—subject always to the wellformation constraints.

19.2.2 Why we should think of NF constructively

The reasons can be roughly grouped as follows

1. Holmes' realizability *aperçu* (see section

   NFHolmes on Realizability for *i*NFonconstructive nature of the proof of the Axiom of Infinity. The interpretation of Peano arithmetic in NF is not at all robust. It does not give rise to an interpretation of Heyting arithmetic into *i*NF.

2. Tie-ups between stratification, normalisability and cut-elimination.

3. No negative interpretation.

4. unstratified nature of the proof of the unsolvability of the halting problem.

I think the first work on intuitionistic NF (which the NF istes have decided to call *i*NF) and its associated type theories was in my Ph.D. thesis. The first *serious* work is much more recent, in Dzierzgowski's Ph.D. thesis.

The two fundamental theorems about NF that we need are both proved by Specker. The first is that NF is equiconsistent with a version of simply typed set theory (in the style of Ramsey or Russell) equipped with an axiom scheme of what would nowadays be called *polymorphism*. To be precise, the language of simple type theory has $=$ and $\in$ and typed variables, so that '$x_n \in y_{n+1}$' is wellformed but '$x_n \in y_n$' (for example) is not. Similarly '$x_n = y_{n+1}$' is not wellformed but '$x_n = y_n$' is. There are axiom schemes of extensionality and comprehension. If ϕ is an expression in this language then ϕ^+ is the result of raising all type indices in ϕ by 1. NF is then equiconsistent with simple type theory plus the axiom scheme $\phi \longleftrightarrow \phi^+$: the scheme of **typical ambiguity**.

Much vaguer is the parallel between the completeness theorem for stratified formulæ in terms of permutation-invariance (offprint attached) and the Läuchli-realizability completeness theorem for intuitionistic logic.

Specker [ ] proved that NF is equiconsistent with a version of simply typed set theory (in the style of Ramsey or Russell) equipped with an axiom scheme of what would nowadays be called polymorphism.

It might seem to the reader that the axiom scheme of typical ambiguity ought to be provable. If this seems obvious it is probably because the reader is confusing this with something that looks rather similar but genuinely *is* obvious. Since ϕ is an axiom iff ϕ^+ is an axiom, we can infer that ϕ is a theorem iff ϕ^+ is a theorem. This is not the same as saying that $\phi \longleftrightarrow \phi^+$ is a theorem! Specker gives illustrations of counterexamples in geometry in [ ]. (Commentary provided: the original article is in German! now translated in Garland, collected metaQuine: a version cleansed of typos is available on my home page))

However, as Randall Holmes has pointed out, if we are thinking in terms of realizability, the proof that any proof of ϕ can be uniformly transformed into a proof of ϕ^+ becomes a realization of $\phi \longleftrightarrow \phi^+$! This is no use to us classically, but it might well turn out to be useful constructively. Specker's equiconsistency lemma is a classical result of course, but there is a constructive treatment of it, due to Dzierzgowski.

19.3 Some funny inductive definitions

If we are going to extract strong consequences (such as excluded middle or an implementation of $\mathbb{N}$) from *i*NF, we have to think about which distinctive features of *i*NF might do the work. It's not stratified separation, because constructive ZF has that and doesn't prove excluded middle; ditto nonexistence of atoms; it's not the existence of a universal set because TST has that (well, sort-of!)

The thing that is distinctive about NF is that it proves outright the existence of least-fixed-points for operations, by the straightforward device of intersection over all sets closed under whatever-the-operation-is.

So consider the following gadget. The $\subseteq$-least set containing V and closed under $X \mapsto \bigcap\{y : X \subseteq \sim\sim y\}$ and arbitrary intersections. "So what?" you might say: TST can do the same. Yes it can, but in the *i*NF context we have the possibility of saying that the intersection of this set is equal to its own power set ... or something else that you can't do in TST. Worth a try.

Let's start by having a look at $\bigcap\{y : \sim\sim y = V\}$. Let's call this thing V_1. Do we have $\sim\sim V_1 = V$? I hope not! $\neg\neg$ distributes over $\wedge$ so $\sim\sim x = \sim\sim y = V. \rightarrow .\ \sim\sim(x \cap y) = V$ but we need to check the $\forall$ case. $x \in \sim\sim V_1$ is

$$(\forall y)((\forall z)(\neg\neg(z \in y)) \rightarrow x \in y)$$

So $\neg\neg(x \in \sim\sim V_1)$ is

$$\neg\neg(\forall y)((\forall z)(\neg\neg(z \in y)) \rightarrow x \in y)$$

so $\sim\sim V_1 = V$ would be

$$(\forall x)\neg\neg(\forall y)((\forall z)(\neg\neg(z \in y)) \rightarrow x \in y))$$

How about $\sim\sim V_1 = V_1$? Starting from So $\neg\neg(x \in \sim\sim V_1)$, which is $\neg\neg(\forall y)((\forall z)(\neg\neg(z \in y)) \rightarrow x \in y)$ we can push '$\neg\neg$' inside '$\forall$'...

$$(\forall y)\neg\neg((\forall z)(\neg\neg(z \in y)) \rightarrow x \in y)$$

and, altho' there are various manipulations we can do, they all involve pasting a '$\neg\neg$' in front of '$x \in y$'. So we can be fairly confident that we can't infer $(\forall y)((\forall z)(\neg\neg(z \in y)) \rightarrow x \in y)$.

So perhaps the family we want is the $\subseteq$-least set containing V and closed under $X \mapsto \underline{\sim\sim}\bigcap\{y : X \subseteq \sim\sim y\}$ and $\underline{\sim\sim}$of arbitrary intersections. (Changes underlined). That way everything is equal to its double complement.

Now let is think about what kind of inductions we can do over this family. Is everything in it $\subseteq$-downward closed? Is everything in it a power set...? That would be nice, beco's then the intersection might be its own power set.

We proved earlier that an arbitrary intersection of stable sets is stable.

This does not make $\{x : x = \sim\sim x\}$ into a basis of closed sets for a topology on V beco's it's not closed under binary unions ($\neg\neg(p \vee q)$ doesn't imply $\neg\neg p \vee \neg\neg q$).

We don't seem to be able to prove that V_1 is nonempty. Perhaps the definition we want is

$$D(X) = \sim\sim\bigcap\{\mathcal{P}(y) : X \subseteq \sim\sim\mathcal{P}(y)\}$$

That way $D(X)$ is always the double complement of a power set (an arbitrary intersection of power sets is a power set, even constructively) and is nonempty. So we'd have to reach a fixed point.

We'll reach a fixed point. What can we say if $X = D(X)$?

$$(\forall z)(z \in X \longleftrightarrow \neg\neg(z \in \bigcap\{\mathcal{P}(y) : X \subseteq \sim\sim\mathcal{P}(y)\}))$$

$$(\forall z)(z \in X \longleftrightarrow \neg\neg((\forall \mathcal{P}(y))(X \subseteq \sim\sim\mathcal{P}(y) \to z \in \mathcal{P}(y))))$$

gulp. That's a lot of work.

Anyway, i don't think there's any way to prevent the process hitting $\{\emptyset\}$ immediately.

Perhaps what we want is: $D(x) = \bigcap\{y : \neg\neg(y = x)\}$. Evidently $D(x) \subseteq x$. I don't think D is $\subseteq$-monotone.

Can we be sure that $D(V)$ is nonempty?

Is the following consistent?

$(\forall z)\neg(\forall y)(\neg\neg(y = V) \to z \in y)$

Consider the inductively defined set $\mathcal{D}$ that is the intersection of all sets closed under directed unions and union-with-disjoint-singletons. (The generalisation of Nfinite).

Evidently every set X in $\mathcal{D}$ is discrete, or whatever we call it: $(\forall x, y \in X)(x = y \vee x \neq y)$.

Is every discrete set in $\mathcal{D}$? How can we put to good use the fact that $\mathcal{D}$ is a set?

Evidently $\mathcal{D}$ is closed under taking bijective copies. Surjective images not so clear

What about the version where we drop the disjointness condition on the singletons? Is there any reason to suppose that the set we obtain is not V?

Can you pull $\neg\neg$ out past an $\exists!$?

Suppose $(\exists! x)\phi(x)$. This is

$(\exists x)(\neg\neg\phi(x) \wedge (\forall y)(\neg\neg\phi(y) \to y = x))$.

So: no, beco's we can't pull the $\neg\neg$ past the $\forall$.

I think we can prove by induction on Nfinite sets that if x and y are Nfinite then $|x| = |y| \vee |x| \neq |y|$.

We prove by induction on 'x' that $(\forall y)(|x| = |y| \vee |x| \neq |y|)$. We don't need `succ` to be injective.

Remark 98 *Every Kfinite subset of an Nfinite set is Nfinite.*

Proof: Suppose true for X, and suppose $x \notin X$. We want every kfinite subset of $X \cup \{x\}$ to be Nfinite. We do this by induction. $\emptyset$ is a kfinite subset of $X \cup \{x\}$ and is Nfinite. Now suppose Y is a Kfinite subset of $X \cup \{x\}$ that happens to be Nfinite. Let $Y \cup \{y\}$ be a kfinite subset of $X \cup \{x\}$. Then either $y \in X$ in which case $Y \cup \{y\}$ is Nfinite by induction hypothesis, so $y = x$ but then $Y \cup \{x\}$ is the union of an Nfinite set and a disjoint singleton and so is Nfinite as desired.

$Y \subseteq X$ is X-stable if $(\forall x \in X)(\neg\neg x \in Y \to x \in Y)$. If X is Nfinite then all its X-stable subsets are Nfinite.

If X is Nfinite then $X \setminus \{a\}$ is X-stable and therefore Nfinite

$X \setminus \{a\} = X \setminus \{b\} \to \neg\neg(a = b)$

Want: the set of Nfinite subsets of an Nfinite set is Nfinite

Another thing we will need. Suppose X is Nfinite. then

$(\forall Y \in \mathcal{P}_{kfin}(X))(\forall y \in X)(y \in Y \vee y \notin Y)$.

Obviously we prove this by induction. Suppose true for X. Consider $X \cup \{x\}$, with $x \notin X$. Now we prove by induction on Kfinite Y that $(\forall y \in X \cup \{x\})(y \in Y \vee y \notin Y)$. True for $Y = \emptyset$. Now suppose it true for Y, and take $z \in X \setminus Y$. Let y be an arbitrary member of X. We want $y \in (Y \cup \{z\}) \vee y \notin (Y \cup \{z\})$. By induction hypothesis we have $y \in Y \vee y \notin Y$. If $y \in Y$ we are OK, so consider the other horn. Since $y \in Y$ and $z \in X$ we have $y = z \vee y \neq z$. If $y = z$ then $y \in Y \cup \{z\}$ and not otherwise. ■

General idea. If we have a map f from V to a Kfinite set X, then we have $(\forall y_1, y_2 \in X)(\neg\neg(y_1 = y_2 \vee y_1 \neq y_2))$, and so, by Johnstone-Linton, $\neg\neg(\forall y_1, y_2 \in X)(y_1 = y_2 \vee y_1 \neq y_2)$. In particular we will have $\neg\neg(\forall x_1, x_2)(f(x_1) = f(x_2) \vee f(x_1) \neq f(x_2))$. Next we can consistently drop the $\neg\neg$ to obtain $(\forall x_1, x_2)(f(x_1) = f(x_2) \vee f(x_1) \neq f(x_2))$, which implies

$(\forall x_1, x_2)(f(x_1) = f(x_2) \vee x_1 \neq x_2)$.

and we hope that $f(x_1) = f(x_2)$ tells us something sensible about x_1 and x_2.

Suppose $\mathcal{V}$ exists. Consider the function $x \mapsto (\sim\sim x) \cap \mathcal{V}$. Call this $K(x)$ for short. Fix an x for the moment. We note that $(\forall y \in \mathcal{V})\neg\neg(y \in x \vee y \notin x)$. By Johnstone-Linton we infer $\neg\neg(\forall y \in \mathcal{V})(y \in x \vee y \notin x)$. From here on we "argue inside the not-nots". Next we prove by induction that if X is an Nfinite subset of $\mathcal{V}$ then $X \cap (\sim\sim x)$ is Nfinite. True for the empty set. Suppose true for X. Consider $X \cup \{z\}$ with $z \in \mathcal{V} \setminus X$. $(X \cup \{z\}) \cap (\sim\sim x)$ is $(X \cap (\sim\sim x)) \cup (\{z\} \cap (\sim\sim x))$. $X \cap (\sim\sim x)$ is Nfinite by induction hypothesis, and—since $z \in \mathcal{V}$ we know we must have $z \in x \vee z \notin x$—so the other term is either the empty set or is a disjoint singleton. Either way $(X \cup \{z\}) \cap (\sim\sim x)$ is Nfinite.

Since $\mathcal{V}$ itself is an Nfinite subset of $\mathcal{V}$ we have proved that, for any x, $\mathcal{V} \cap (\sim\sim x)$ is not-not Nfinite. So in particular, for any term t, we have $\mathcal{V} \cap (\sim\sim t)$ is not-not-Nfinite so we can safely assume that $\mathcal{V} \cap (\sim\sim t)$ is Nfinite. So we have safely added the scheme that $K(t)$ is Nfinite. Now let t_1 and t_2 be two closed terms. $K(t_1)$ and $K(t_2)$ both belong to $\mathcal{P}_{Kfin}(\mathcal{V})$, the set of Kfinite subsets of $\mathcal{V}$. We have $(\forall x, y \in \mathcal{P}_{Kfin}(\mathcal{V}) \neg\neg(x = y \vee x \neq y)$. Now $\mathcal{P}_{Kfin}(\mathcal{V})$ is a Kfinite set, being the set of Kfinite subsets of a Kfinite set. So we can use Johnstone-Linton to infer $\neg\neg(\forall x, y \in \mathcal{P}_{Kfin}(\mathcal{V})(x = y \vee x \neq y)$.

In particular we get

$\neg\neg(K(t_1) = K(t_2) \vee K(t_1) \neq K(t_2))$

Now whenever we have a proof of $\neg\neg p$ (with p closed) we can consistently assume p so we can drop the $\neg\neg$ to assume

$$K(t_1) = K(t_2) \vee K(t_1) \neq K(t_2)$$

If $K(t_1) \neq K(t_2)$ then $t_1 \neq t_2$. If $K(t_1) = K(t_2)$ then we argue as follows. Suppose

$\neg\neg(x \in t_1)$ iff

$\neg\neg(x \in t_1) \wedge \neg\neg(x \in \mathcal{V})$ iff

$\neg\neg(x \in (\sim\sim t_1)) \wedge x \in \mathcal{V}$ iff

$x \in (\sim\sim t_1) \cap \mathcal{V} = K(t_1)$ iff

Now $K(t_1) = K(t_2)$ so we can retrace our steps...

$x \in ((\sim\sim t_2) \cap \mathcal{V}) = K(t_2)$ iff

$\neg\neg(x \in (\sim\sim t_2)) \wedge x \in \mathcal{V}$ iff

$\neg\neg(x \in t_2) \wedge \neg\neg(x \in \mathcal{V})$ iff

$\neg\neg(x \in t_2)$

So $(\sim\sim t_1) = (\sim\sim t_2)$

We conclude that $t_1 \neq t_2 \vee ((\sim\sim t_1) = (\sim\sim t_2))$

In particular if $t_2 = \sim\sim t_1$ we deduce that we can safely assume that, for any closed term t, either $t = \sim\sim t$ or $t \neq \sim\sim t$

$\{\emptyset\}$ is a wellordering. If $\mathcal{X}$ is a wellordering and $x \notin \bigcup \mathcal{X}$ then $\mathcal{X} \cup \{\bigcup \mathcal{X} \cup \{x\}\}$ is a wellorder. A union of a $\subseteq$-chain of wellorderings is a wellordering. An ordinal is an isomorphism type of wellorderings.

If there is a dense Kfinite set, there is a dense Nfinite set. If X is dense and Kfinite, then the quotient over not-not= is a surjective image of a Kfinite set and so is Kfinite. Now it is in any case discrete (being a partition) so it is Nfinite. Nfinite sets have transversals. This transversal is the dense Nfinite set we are after..

(i) Commutation of $\neg\neg$

We appeal to clause 4 of remark 64.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\forall x\neg\neg\phi(x)}{\neg\neg\phi(x)}\,\forall\text{-elim}\qquad [\neg\phi(x)]^1}{\perp}\,\to\text{-elim}}{\phi(x)}\qquad [\phi(x)]^1 \qquad \phi(x)\vee\neg\phi(x)}{\phi(x)}\,\vee\text{-elim}}{\forall x\phi(x)}\,\forall\text{-int} \tag{10}$$

This proves something slightly stronger than commutation. There is a converse: commutation implies double negation for stratified formulæ:

and we know from page 57 that double negation for atomics implies excluded middle for atomics, thereby closing the circle. ■

Dear Dr Drago,

Please forgive me writing to you out of the blue like this, but your name came up in a google search on the above topic. I have a medium-term project to understand the constructive version of Quine's NF, and in the course of this i am moved to investigate the symmetric group on the universe, which is of course a set in NF. I have a feeling that constructive NF ought not to be able to prove the existence of any nontrivial permuations unless excluded middle holds - or at least that there is a weaker result of that nature to be had. This has caused me to consider the group of permutations that are not-not equal to the identity, which is of course a normal subgroup of Symm(V).

But less of that! Is there a good place to start reading about constructive group theory?

yours

Thomas Forster

Dear Thomas Forster,

my work as historian of Physics included very lttle about constructive group theory. Anyway the reference text is by A course in constructive algebra by Ray Mines, Fred Richman, Wim Ruitenburg -Springer 1988. They follow an idea which I do not share: the definition of the inverse element by means of the notion of apartness, which essentially includes Markov principle, transcending constructivism.I think that a constructive theory of group is the crucial knot of the research on applied mathematics.

But I was very surprised to read not-not equal to the identity; never I met this expression in my studies on group theory; where you found out this definition? What means it in mathematical terms? In my past work I obtained evidence for the great importance of the double negated statements which are not equivalent to the correspondig affirmative ones (in this case: identity).

Thanks for your answer

All the best

Antonino Drago

But less of that! Is there a good place to start reading about constructive group theory?

Contact Fred Richman ¡richman@fau.edu¿.

GS

There's also "A Course in Constructive Algebra," by Mines, Richman, and Ruitenburg, Springer '88.

Bob Lubarsky

—–Original Message—–

From: fom-bounces@cs.nyu.edu [mailto:fom-bounces@cs.nyu.edu] On Behalf Of Andrej Bauer

Sent: Saturday, April 02, 2011 2:45 AM

To: Foundations of Mathematics

Cc: T.Forster@dpmms.cam.ac.uk

Subject: Re: [FOM] Constructive Group Theory

But less of that! Is there a good place to start reading about constructive group theory?

A place to start is the second volume of Troelstra and van Dalen's "Constructivism in mathematics". There must be other sources though, which cover more.

With kind regards,

Andrej

A couple of basic principles for constructive set theory ...

$$(\forall xy)(\neg\neg(x \in y) \to (\exists y')(\neg\neg(y = y') \wedge x \in y'))$$

$$(\forall xy)(\neg\neg(x \in y) \to (\exists x')(\neg\neg(x = x') \wedge x' \in y))$$

The second one implies excluded middle. For consider: $\neg\neg(y \in \{z : z = y \wedge (p \vee \neg p)\})$. Our principle would imply $(\exists x)(\neg\neg(y = x) \wedge x \in \{z : z = y \wedge (p \vee \neg p)\})$. So $\{z : z = y \wedge (p \vee \neg p)\}$ is inhabited. So $p \vee \neg p$.

That trick will not work on

$$(\forall xy)(\neg\neg(x \in y) \to (\exists y')(\neg\neg(y = y') \wedge x \in y'))$$

$$(\forall xyy')(x \in y \wedge \neg\neg(y = y') \to (\exists x')(\neg\neg(x = x') \wedge x \in y'))$$

$$(\forall x y y')(\neg\neg(x \in y) \wedge \neg\neg(y = y') \to (\exists x')(\neg\neg(x = x') \wedge x \in y'))$$

$\models (\forall xy)((\forall z)(z \in x \longleftrightarrow z \in y) \to x = y)$

$(\forall W)(\forall xy)(W \models (\forall z)(z \in x \longleftrightarrow z \in y) \to x = y)$

$(\forall W)(\forall xy)(\forall W' \geq W)([W' \models (\forall z)(z \in x \longleftrightarrow z \in y)] \to W' \models x = y)$

$(\forall W)(\forall xy)(\forall W' \geq W)([(\forall W'' \geq W')(\forall z)(W'' \models (z \in x \longleftrightarrow z \in y))] \to W' \models x = y)$

$(\forall W)(\forall xy)(\forall W' \geq W)([(\forall W'' \geq W')(\forall z)(\forall W''' \geq W'')((W''' \models z \in x \longleftrightarrow W''' \models z \in y)] \to W' \models x = y))$

How about cooking up a model of *i*NF by making use of the fact that the nasty bits of recursion theory depend on punning? Taking Kripke conditions to be enumerations of total recursive functions satisfying extensionality or something . . . (jan 1997)

Notice that the fixed point theorem is stratified in this sense but presumably has no stratified proof.

There is this idea that *i*NF should be consistent, and that we should be able to prove the consistency inside a pretty small part of arithmetic by reasoning about recursive functions. There are two major problems. One is the axiom of complementation: the complement of an r.e. set is not an r.e. set. One bright idea i had was that the complement of a function should be a function that accepts n and returns a function that behaves like a complement of that function for the first n steps, but then the complement of a complement of f would be K of f and stratification goes out of the window.

. . . and the second is extensionality. Rice's theorem will tell us that for any turing machine there is a non-recursive set of machines that have exactly the same behaviour not only in the sense that they produce the same answers but that they take the same length of time to do it. This means that we will have to take our numbers to be functions f that take two arguments, i and t, and return the state of the universal turing machine with input f and i and return its state after t steps of computation.

Or . . . thinking aloud . . . seek the least fixed point for the operation that accepts an equivalence relation on naturals and returns the equivalence relation that sez two functions are similar if they send similar inputs to **identical** outputs. The least fixed point is a PER not an equivalence relation, since if $1 \sim 2$ any function f s.t. $f`1 \neq f`2$ cannot resemble anything!

Index possible worlds by $\mathbb{N}$. Say $n \models f \in g$ iff $f`g \downarrow \leq n$. That will give us $\neg\neg p \vee \neg p$ for atomic p. I don't *think* that's too strong.

The obvious way to prove Con(*i*NF) is to use recursive functions with M thinks that f ∈ g if f halts on g and gives one in fewer than m steps, or something like that. The we use a realisability trick (like M thinks that $\forall x \phi(x, y)$ if for all M' > M, and for all x, M' thinks that $\phi(x, y`m')$.) The problem with this whole approach is explaining why this doesn't give us the Russell class. there is no obvious reason why stratification helps.

(I suppose the answer to that question is this: think of programs not functions. That was obvious wasn't it! But the point is that you can't tell by looking at the program for the Russell class that it is the program for the Russell class. It's something to do with properties being Δ_0)

(Is the way into understanding Realizability the infinite regress that one is launched on by the problem of the nonconstructive proof that the range of any nondecreasing total computable function is decidable?)

The trouble with trying to make use of the insight that it's only unstratified stuff that enables us to prove the unsolvability of the halting problem is that a perfectly respectable piece of innocent code might just happen to code the self-application function—for a gnumbering of programs that we just hadn't spotted. This means that we have to index our worlds by gnumberings and ensure that we don't allow as individuals of any world any functions which satisfy naughty things. This will have the effect that different worlds have different individuals and this makes the logic of quantifiers nasty (i'm so used to hanging around modal logicians that the first thing that comes to mind is the Barcan formula and its converse, both of which fail in these conditions). Perhaps the trick is to have *only* nonrecursive gnumberings!

$H(f, n, k)$ notational variant of: $f(n)\downarrow = k$, where the 'n' can be a tuple.

We have term-forming operators like $\circ$ and rules like $\dfrac{H(f, i, j) \quad H(g, j, k)}{H(g \circ f, i, k)}$.

We will also have a term-forming operator $PR(f, g)$, which denotes the function declared by primitive recursion over f and g with rules

$$\frac{H(f, \vec{v}, k)}{H(PR(f, g), 0 \frown \vec{v}, k)}$$

$$\frac{H(PR(f, g), n \frown \vec{v}, x) \quad H(g, \langle x, n+1, \vec{v}\rangle, k)}{H(PR(f, g), (n+1) \frown \vec{v}, k)}$$

and a similar rule for μ-recursion.

want to fail to refute something like this:

$(\forall fn)[H(A(f), n, 0) \longleftrightarrow (\exists k)(H(f, n, k)) \wedge H(A(f), n, 1) \longleftrightarrow (\forall k)(\neg H(f, n, k))]$

for each term $A(f)$ with f free.

Or do we want variables of all types? So that A is not a context but a variable? The we need higher-type operations like primitive recursion, composition and so on.

So how about pointed sets with extensional relations on them ("What about pointed sticks?!") but this doesn't work by itself, because we cannot get a universal set. Why not? Consider countable widgets (these things are called widgets for the moment). There are uncountably many countable ordinals. We might be able to do something with recursive widgets.... One also has the feeling that somehow the fact that there are universal turing machines ought to help....

Notice that the unsolvability of the halting problem is a stratified assertion! But it doesn't have a stratified proof! Or does it ...? fri 6/iii/98

Lower case roman letters are in $\mathbb{N}$ or are TM's: same thing.

Say m **simulates** n if there is $k \in \mathbb{N}$ such that for all $j \in \mathbb{N}$, $m(\langle j, k\rangle) \sim n(j)$ where $\sim$ means halts on the same inputs and gives the same outputs.

The set of gnumbers of partial functions whose values lie in $\{0,1\}$ is not decidable. So this is what we do. The problem all along has been to prevent there being accidentally a function that happens to be self-application. Take the domain to be the set of partial functions whose values lie in $\{0,1\}$, with a (necessarily non-recursive) enumeration. That also takes care of extensionality.

So a possible world is a recursive set W of gnumbers of partial recursive fns $\mathbb{N} \to \{0,1\}$, and for each (graph of a) total recursive function $\mathbb{N}\to \{0,1\}$ W contains a gnumber of total recursive function $\mathbb{N} \to \{0,1\}$ with the same graph. Each such set can see all its subsets. Alternatively we can think of the sets as squashed, so that numbers do not represent the same functions in all poss worlds, except asymptotically. That is to say, for every number n, there is a total recursive function $\mathbb{N} \to \{0,1\}$ that it notnot codes.

Fix X a selection set for the family of equivalence classes of gnumbers of total recursive functions $\mathbb{N} \to \{0,1\}$ under the relation of having the same graph. No such selection set can be decidable, which is good. Possible worlds will be recursive supersets

`From holmes@catseye.idbsu.edu Tue Jan 19 23:28:15 1999`

There is a subtlety about the definition of "classically inductive": it is unclear how to define inductive set in SF, where the union $x \cup \{a\}$ might not be uniquely determined. I imagine the correct definition would be for an inductive set to contain all unions of its elements with singletons (all candidates for being $x \cup \{a\}$); but this really does not matter in this proof, because the Kfinitude of V implies that there really is only one candidate for being $x \cup \{a\}$ (up to the double complement of equality).

Randall

`From holmes@catseye.idbsu.edu Thu Jan 21 16:20:53 1999`

The conclusion to be drawn is that the negative interpretation of classical NFU remains interesting, as does the fact that Ω Kfinite is strong, but the result that V is not Kfinite is too easy.

–R.

`From holmes@catseye.idbsu.edu Thu Jan 21 16:48:12 1999`

*i*NF + commutation proves that sets with the same double complement are not not equal, so the double negation interpretation of classical SF obtained from it is extensional: thus this system interprets classical NF and of course interprets constructive arithmetic (because it interprets classical arithmetic!)

–Randall

`From holmes@catseye.idbsu.edu Thu Jan 21 16:50:31 1999`

The only caveat being that you might have some special sense in mind of "interprets constructive arithmetic"? I have no idea whether Dzierzgowski's favorite form of Infinity holds, for example; I just have a classical interpretation of arithmetic inherited from the embedded classical NF.

–Randall

`From holmes@catseye.idbsu.edu Thu Jan 21 16:52:38 1999`

Precisely. The interesting models of *i*NF (if there are any) have infinite truth value algebras. –Randall

From Daniel (forwarded by Randall)

Yes, you can prove that if x is Kfinite then not not (x is Nfinite). This is a consequence of Remark 4.3 of my paper in Notre Dame Journal (vol 17, no 4, 1996, 585-601). Make sure this makes it into the biblio

Details of the proof are not given in the paper. I think that you can prove by induction on x that if x is Kfinite then not not exists y Nfinite s.t. x is a subset of y. Then prove by induction on y that if y is Nfinite and y' is any subset of y, then not not y' is Nfinite. Then reduce not not not not to not not and you've got it.

But you cannot prove that if x is Kfinite then there exists y such that y is Nfinite and not not ($x = y$). I have some unpublished properties of finite sets. I think I sent you some draft notes about this some time ago, didn't I? As I do not remember myself if I did, I won't blame you if you don't remember either!

Best wishes,

Daniel.

`From ddz@skynet.be Mon Jan 25 21:40:13 1999`

Thomas,

I'm happy to see again some interest in *i*NF!

Hum... I remember I wrote you some messages about negative interpretation quite a long time ago. But I don't remember I have shown that if you add excluded middle for equality, you have a negative interpretation of NF... A fortiori, I don't know about $x \neg\neg = y \vee x \neq y$. I'll check my own notes. But, as I told Randall, I'm very busy until next week. I can't think about all this now.

I do think that HA can be interpreted in *i*NF + (the right) infinity. I had a hint for proving this, by finding the right topos. The idea was quite technically difficult to handle but I think it could work.

I'll come back next week.

Best wishes,

Daniel.

`From holmes@catseye.idbsu.edu Wed Jan 27 20:04:23 1999`

Dear Thomas and Daniel,

This is my program (even more improved version).

1. V is not Nfinite. If it were, it would be discrete, and from this we could deduce excluded middle for weakly stratified formulas, so that we would be able to prove stratified sentences of classical NF such as "V is not Nfinite".

2. Since V is not Nfinite, any Nfinite set x is not equal to V, and so it is not not the case that there is a z which is not an element of x.

3. From this it follows that if an Nfinite cardinal n is inhabited, n+1 is not uninhabited, from which it follows by induction that all Nfinite cardinals are not uninhabited.

4. Prove that any inhabited Nfinite cardinal m+1 has unique predecessor (if it is of the form $x \cup \{y\}, x$ disjoint from $\{y\}$ and Nfinite, x is of size m). This is true for 0, vacuously. Suppose it true for m – we consider $x \cup \{y\}, x \in m, x$ and $\{y\}$ disjoint, and suppose that $x \cup \{y\} = x' \cup \{y'\}$. Either $y' \in x$, thus $y' \neq y$ or $y' \in \{y\}$, thus $y' = y$. If $y' = y$, we can conclude $x' = x$ (can we–yes, by discreteness of Nfinite sets?), so $x' \in m$. Otherwise $y' \in x$, so x is inhabited and m has a uniquely determined predecessor m-1 by ind hyp. Since x is Nfinite, it is discrete, and is equal to $(x - \{y'\}) \cup \{y'\}$, and by induction $(x - \{y'\})$ is of cardinality m-1, and $(x - \{y'\}) \cup \{y\}$ belongs to m as required.

4. Prove by induction that if n is an inhabited Nfinite cardinal, any set x belongs to n iff there is a bijection between x and some (thus all) elements of n.

5. One proves by a similar induction that no Nfinite cardinal has an element with a bijection between it and a proper subset. if it did, a set with cardinality one less would. Unique predecessor is needed for this.

6. Prove that for all Nfinite cardinals, $n \neq n+1$. One also needs that $n+1 \notin \{1...n\}.\{1...n\}$ = the set of cardinals of inhabited Nfinite subsets of elements of n. n+1 cannot belong to this by absence of bijections to proper subsets. (this also handles n+1 $\neq$ n). This relies on the not uninhabited nature of all Nfinite cardinals.

7. Prove by induction that each Nfinite cardinal m is inhabited by some initial segment $\{1...n\}$ of the Nfinite cardinals. This relies on $n+1$ being a fresh object at each stage.

8. Thus every Nfinite cardinal is inhabited.

This is the Axiom of Infinity in the form required for *i*NF.

–Randall

Later (tf): the mistake here is that $\neg\forall$ does not imply $\neg\neg\exists\neg$.

`From holmes@catseye.idbsu.edu Wed Jan 27 21:02:53 1999`

I would need to show that $\{1...n+1\} = \{1...n\} \cup \{n+1\}$, which is not glaringly obvious but might be true. –Randall

`From holmes@catseye.idbsu.edu Wed Jan 27 21:14:30 1999`

This needs another detail. It is not only necessary to show that $n+1$ is not in $\{1...n\}$, but it is also necessary to show that $\{1...n\} \cup n+1 = \{1...n+1\}$. I think that it is possible to show this, but it involves work!

0 is inhabited by $\{1...0\} = \emptyset$. If n is inhabited by $\{1...m\}$, then n+1 is inhabited by $\{1...m+1\} = \{1...m\} \cup \{m+1\}$ where I can definitely show $m+1 \notin \{1...m\}$ but am less sure about the equation $\{1...m+1\} = \{1...m\} \cup \{m+1\}$.

–Randall

`From holmes@catseye.idbsu.edu Wed Jan 27 21:36:58 1999`

Jottings in support of my "program" (mainly of the missing link $\{1...n\} \cup \{n+1\} = \{1...n\}$). The issue is the well-definedness of subtraction for Nfinite cardinals, which I believe I establish.

Prove by induction that an Nfinite set minus a singleton is Nfinite of a uniquely determined cardinality:

vacuously true of 0!

Suppose true for sets of size n. Let A be of size n and consider $A \cup \{x\}$, $x \notin A$. Remove an element y from this set. y is either equal to x, in which case we obtain the set A and we succeed, or it belongs to A, in which case we obtain $(A - \{y\}) \cup \{x\}$, where $A - \{y\}$ is Nfinite by inductive hypothesis and also of uniquely determined cardinality by ind hyp.

I need to show that any Nfinite subset of an Nfinite set has Nfinite relative complement.

Certainly true if the smaller set has size 0.

Suppose the smaller set has size $n + 1$. Call the smaller set A and the larger set B. $A = C \cup \{x\}$ for some C and x. C is an Nfinite set of size n. The complement of C is Nfinite by ind hyp, and the complement of $C \cup \{x\}$ is obtained by removing x from this set, and so is also Nfinite (by the result that removing one element from an Nfinite set leaves an Nfinite set).

Now can we show that any Nfinite subset of a set of size $n + 1$ is either of size $n + 1$ or is an Nfinite subset of a set of size n?

It is either empty or inhabited. If it is inhabited, it is a subset of the set of size n obtained by deleting the inhabitant. Thus $\{1...n+1\} = \{1...n\} \cup \{n+1\}$.

–Randall

`From holmes@catseye.idbsu.edu Thu Jan 28 15:29:28 1999`

$n + 1$ is always defined! I define $n + 1$ as the set of all disjoint unions of elements of n with singletons. In the classical case, the last natural number might be empty; showing that all natural numbers are not uninhabited shows (partially) that this doesn't happen.

–Randall

`From t.forster@dpmms.cam.ac.uk Thu Jan 28 15:36:49 1999`

Yes, i think i see what you mean. We can prove by induction that every Nfinite cardinal is nonempty can't we. Let me see...

Suppose n is nonempty. If a is an arbitrary thing in n, then there is not-not something in $-a$. If there were such a w, then $a \cup \{w\}$ would be in $n + 1$. But if $n + 1$ is empty there is no such w, but there not-not is such a w so there is not-not something in $n + 1$—as long as there is something in n. But n is nonempty.

Looks OK to me.

Well done!!

`From holmes@catseye.idbsu.edu Thu Jan 28 17:25:11 1999`

To convert any negative weakly stratified sentence into a negative atomic formula:

$\neg\phi = x \in \{x|\neg\phi\}$ ('x' not free in ϕ) which is equivalent to $\neg\neg x \in \{x|\neg\phi\}$

so we have $\neg x \in \{x|\neg\phi\}$ or not not $x \in \{x|\neg\phi\}$

which is equivalent to

$x \in \{x|\neg\phi\}$ or $\neg x \in \{x|\neg\phi\}$

which is equivalent to

$\neg\phi \vee \neg\neg\phi$

for any weakly stratified ϕ.

This is clearly enough to interpret NF.

`From holmes@catseye.idbsu.edu Thu Jan 28 18:13:59 1999`

Now I think I see an example.

Let type 0 of a model of ITT contain a single object 1.

There are ω stages of knowledge.

1 has approximations to which 1 is first seen to belong at each of the omega stages. Each of these objects is not not equal to 1.

The set $\{\{1\}\}$ and its double complement are frankly unequal! The problem is that each of the ω approximations to $\{1\}$ belong to the double complement at every stage of knowledge, but are seen to belong to $\{\{1\}\}$ at different stages; one never sees that each element of the double complement belongs to $\{\{1\}\}$, so this isn't true!

The double complement of $\{\{1\}\}$ is not Nfinite, because an Nfinite set either has 0 elements, 1 element, or at least 2 distinct elements, and none of these conditions holds of the double complement.

How sickening :-)

–Randall

`From holmes@catseye.idbsu.edu Thu Jan 28 18:19:31 1999`

The elegant way to interpret classical SF in *i*NF + there is an Nfinite set which is dense is to assign each object its intuitionistic complement as its classical extension. It is then obvious that comprehension holds, that stratified reasoning is classical, and that there is no reason to believe that extensionality holds – so far...

–randall PS I thought you would like that.

`From holmes@catseye.idbsu.edu Thu Jan 28 18:20:52 1999`

Something needs to be done with equality in that picture – but one could always recall that classical SF without equality allows a definition of equality...

`From holmes@catseye.idbsu.edu Thu Jan 28 19:43:28 1999`

Here's why I still don't believe that *i*NF is weak:

1. Suppose that all Nfinite cardinals are not uninhabited. The rest of my argument works.

2. Suppose that Ω is Kfinite. We know what happens then.

3. This leaves us with a picture of weak *i*NF in which there is an Nfinite dense set and an infinite Ω. In terms of Kripke semantics, this suggests that we have a universe inhabited by a discrete finite set of objects and various things which are not not one of these objects.

The difficulty is that taking enough power sets of a collection like this with an infinite truth value algebra ought to generate infinitely many distinct objects.

This strongly suggests to me that we will be able to prove Infinity...

–Randall

`From holmes@catseye.idbsu.edu Thu Jan 28 23:32:00 1999`

Dear Daniel,

At this point we don't have a proof of Infinity. I was able to find out for myself that Nfinite sets can have non-Nfinite double complements (thanks to my education from your ITT paper!)

I do know, I think, that if every Nfinite cardinal is not uninhabited, then Infinity holds (each Nfinite cardinal is actually inhabited).

The interesting possibility is that there is an Nfinite set whose double complement is V. If I can show that this is not possible, then I show that each Nfinite set has nonempty complement (this means not uninhabited rather than inhabited), from which the rest of my proof of Infinity would go forward.

So I'm thinking about consequences of an Nfinite set dense in V. One consequence is excluded middle for negations of atomic formulas, from which one can get an interpretation of classical SF.

It appears that the Nfinite set dense in V requires Ω to be not Kfinite (because the interpreted classical SF does not satisfy Infinity); my suspicion is that it may prove (in ITT) that if Ω is not Kfinite and if some type has an Nfinite dense set, some higher type will turn out to have infinitely many distinct elements; this would kill Nfinite dense sets in *i*NF and make a proof of AxInf possible. But I could be quite wrong!

Watch this space!

–Randall

`From holmes@catseye.idbsu.edu Thu Jan 28 23:55:02 1999`

I think that it is reasonably clear that if there is an Nfinite dense subset of V, Ω cannot be Kfinite. The reason why this should be true is that the classical interpretation of SF will tell us that the universe is finite, and so will be nonextensional, from which it follows that forall does not commute with not not, from which it follows that Ω is not Kfinite.

We have the following table:

1. Ω is Kfinite. The classical interpretation of SF yields classical NF and thus infinity.

2. Each Nfinite set has nonempty complement. In this case I still think I can prove Infinity.

3. There is an Nfinite dense subset of V. In this case classical SF with finite universe is interpreted, so Ω cannot be Kfinite.

I don't know whether these alternatives are in any sense exhaustive. Only 3 leaves weakness open as an option, so that's what to study – also, if 3 can be refuted I believe that Infinity then becomes provable– if there is no Nfinite set dense in V, then every nfinite set has a complement which is not not inhabited, and my proof in case 2 goes forward.

Nfinite dense subsets of V are where the action is!!!

–randall

`From holmes@catseye.idbsu.edu Mon Feb 01 17:15:44 1999`

What I'm hoping to do is show in ITT that if Ω is not Kfinite and some type has a finite discrete dense subset, then some higher type is frankly infinite. I don't see how to do this yet, but the semantics strongly suggests to me that this ought to be true. If it isn't, I will of course be sadder but wiser; if this does work, then *i*NF is strong.

–Randall

`From holmes@catseye.idbsu.edu Thu Feb 04 17:30:39 1999`

I think that what we know is this. If we want to investigate the possibility of *i*NF being weak, we might as well assume the following things:

1. The intersection of all dense subsets of a singleton is empty.

   For if this is not the case, the intersection of all dense subsets of a singleton will be dense, $\neg\neg$ will commute with $\forall$, and the interpreted classical SF will be NF, so we are out of the "weak" realm.

2. There is an Nfinite dense subset of the universe.

Suppose that this is not the case. It follows that my proof that the set of Nfinite numerals is infinite goes through, and we are out of the realm of weak theories again, though not in NF proper.

–Randall

`From t.forster@dpmms.cam.ac.uk Sun Feb 07 18:23:39 1999`

A stronger version of inequality gives us a stronger version of finiteness:

$(\forall x)(x \neq a \vee x \neq b)$

If we have a dense finite set all of whose members are distinct in that sense then my argument works. Prove by induction on such finite sets that if you are not-not in it then it has a member not-not equal to you. Then if there is a dense one my argument using surjections works, and one can even prove:

for all x and y either not-not $x = y$ or x and y are strongly unequal as above.

Not sure if this is any use.....

`From t.forster@dpmms.cam.ac.uk Mon Feb 08 16:35:47 1999`

Couldn't sleep last night. I lay awake thinking about the lfp for the operation that takes the relation R to the relation

$\lambda R.\neg(\forall z)(\neg zRa \vee \neg zRb)$

This lpf is of course a set in *i*NF!! Is it any use? Might there be some point in considering sets which are Nfinite in the stronger sense that one can only add elements which are (in this sense) utterly unlike what is already in the set?

The definability of such fixed points is a nice feature of *i*NF. I think we should try to make it work for us somehow.

`From holmes@catseye.idbsu.edu Mon Feb 08 18:01:46 1999`

I think that the existence of objects which are strongly distinct in your sense is a powerful constraint on what the truth value algebra is like. Suppose that any element of the truth value algebra has two stronger and incompatible elements; then **no** pair of objects can be distinct in your strong sense (because at any stage of knowledge I can present a name which is either a name of a or a name of b (speaking classically) but we cannot decide which right now).

–Randall

PS so the truth value algebra needs to be eventually "linear" in some sense for us to make use of this.

`From holmes@catseye.idbsu.edu Mon Feb 08 18:05:49 1999`

I think that the same consequences for the truth value algebra follow if any pair of objects must be either not equal or not not equal.

There is no reason whatever why an Nfinite dense set should have such effects on the truth value algebra.

(or any reasonable kind of finite set).

–Randall

`From holmes@catseye.idbsu.edu Wed Feb 03 23:42:46 1999`

Dear Thomas,

Just some insubstantial musings...

The notion of finitude in the interpreted classical SF applies to some sets which are not Nfinite or Kfinite (under reasonable assumptions).

Suppose that the intersection of all dense truth values is the empty set (if I don't assume this, I have interpreted classical NF, so I might as well :-)) Then the double complement of a double singleton $\{\{x\}\}$ is not Nfinite. For it contains each set $\{y|y = x \wedge d\}$, where d is a sentence with dense truth value. The assertion that $\{\{x\}\}^{c^c} = \{\{x\}\}$ is then at least as strong as the conjunction of all the dense truth values, and this is known to be false.

This means that $\{\{x\}\}^{c^c}$ is not Nfinite. If it were Nfinite, it would either have no elements (but it has $\{x\}$), have exactly one element (but then it would be $\{\{x\}\}$, which it isn't) or have at least two distinct elements (it doesn't!). So it is not Nfinite. But it is certainly finite in the sense of the interpreted classical SF – in the interpreted classical terms, it has exactly one element: any two things which are not not in it are not not equal to each other (because they are both not not equal to $\{x\}$). Moreover, it is likely to be treated as a set in an interpretation of NFU: the natural way to convert the classical SF to classical NFU is to treat double complements as sets and things which are apart from their double complements as nonsets.

I'm planning to think about what the system *i*NF + "interpreted classical NFU says the universe is finite" looks like. I'd like to see what the relationship is between this system and *i*NF + "there is an Nfinite dense subset". But in order to do this, I need to understand what the classical notion of finitude is doing...

I hope you picked up from my previous note that I am quite doubtful that it is really true that excluded middle for negatomics follows from existence of an Nfinite dense set. I really can't see any reason why each object x in V has to have an associated object $f(x)$ in the dense set such that $\neg\neg(x = f(x))$; in fact, I think it is easy to model the contrary in ITT: there will be many possible objects which are not not in the dense set but which have not settled down to being one or another of its elements!

–Randall

`From holmes@catseye.idbsu.edu Wed Feb 17 18:35:08 1999`

Are you reading it? –Randall

I'm thinking about how much of the Kripke model semantics can be represented internally to ITT.

truth values = subsets of a singleton

possible objects correlate with sets such that any two elements of the set are equal: a possible object at any stage of knowledge can be coded by a "near-singleton" in this sense available "now".

We can't hope to say anything about the Kripke model semantics that isn't true of any cofinal substructure of the Kripke model in a suitable sense—since cofinal substructures will satisfy the same sentences of intuitionistic logic. It would be nice if one could say everything that can be said mod cofinal substructures, but I doubt that this is possible. The biggest obstacle I see is saying sensible things about "branching" in the truth value algebra.

`From holmes@catseye.idbsu.edu Thu Apr 29 20:42:57 1999`

Dear Thomas,

I'm running another process in parallel to everything else I'm doing; I'm thinking about infinity in *i*NF.

If *i*NF does not prove Infinity it must be consistent to have an Nfinite set whose double complement is the universe. This implies by stuff we've already done that the intersection of all dense truth values is the False (i.e., $\forall$ does not commute with not not).

I believe that I can establish that there are infinitely many distinct objects if there is any function which sends dense truth values to stronger dense truth values (i.e., sends each dense subset A of a singleton $\{x\}$ to a subset of A which is also dense in $\{x\}$) and which is distinct from the identity function. The idea is this: let f be such a function and consider the sets $\{f^n(y)|y$ is dense in $\{x\}$ and $x \in y\}$; I believe I see how to show that all these sets are distinct.

Do you see any method in *i*NF with an Nfinite dense subset of the universe (and so Ω not Kfinite) to generate stronger dense truth values from given dense truth values?

I wouldn't blame you if you found these concepts rather obscure...

–Randall

`From holmes@catseye.idbsu.edu Wed May 05 21:27:04 1999`

Dear Thomas (cc Daniel),

Without being able to prove anything, I still suspect that *i*NF must be strong. My reasoning is as follows: the truth value algebra Omega is strongly cantorian; this suggests that any automorphism at work in the model theory of *i*NF must fix all elements of the truth value algebra. But any model of *i*NF either has an infinite truth value algebra or interprets classical NF (because *i*NF with Kfinite truth value algebra interprets classical NF). This means that any automorphism at work in a model of *i*NF must fix all elements of an infinite set. Intuitively, this suggests that a version of *i*NF with a non-Kfinite truth value algebra should be very strong (it seems that it ought to satisfy AxCount!) The reason this argument doesn't translate into a proof is that I don't know enough about the model theory of *i*NF to make it rigorous; this is all based on analogies with the model theory of NF or NFU which may break down for reasons I don't see. It may be that models of *i*NF don't necessarily imply models of ITT with automorphisms at all...

Has anyone thought about the relationships between models of *i*NF and models of ITT with automorphisms? Are there any actual results?

–Randall

`From holmes@catseye.idbsu.edu Tue Jan 26 18:50:21 1999`

The reason that excluded middle for weakly stratified formulas implies interpretability of NF is that one can restrict oneself to stratified formulas in proving stratified formulas of NF. Thus we have a classical version of the stratified theory of NF embedded in *i*NF + excluded middle for weakly stratified formulas, and the stratified theory of NF is just as strong as full NF (being equivalent to TT + Amb).

–Randall

`From holmes@catseye.idbsu.edu Wed May 05 22:51:05 1999`

My intuitive argument for the strength of *i*NF admits a possible counterexample. The same argument suggests that any model of *i*NFU + "there is an Nfinite dense subset of the universe" should have Kfinite truth value algebra. This theory certainly has models (any model of classical NFU with finite universe is a model of this). Does it have models with non-Kfinite truth value algebras?

–Randall

`From holmes@catseye.idbsu.edu Mon Jun 21 17:34:01 1999`

The point being that the natural way to collapse a 3-valued model of NFU to a 2-valued model of NFU has the embarrassing feature that it manufactures lots of urelements which are descendents of *sets* rather than urelements of the old model. This would cease to be an embarrassment if one had some reason to believe that the *new* model had indiscernbile urelements. One needs to show that "the same things are true" in the two models (of course, the situation in the 2-valued model may be clearer, but nothing false in the 3-valued model may be true in the 2-valued model); the creation of urelements from old sets blocks the usual way to prove this.

–Randall

`From t.forster@dpmms.cam.ac.uk Mon Aug 23 10:58:03 1999`

Dear Dr. Bell,

I have just picked up the current number of the JSL and found your article about Frege's theorem. (I'd never heard it called that but even at my age one learns something new every day). This stuff is of great interest to me—and to the two people i am cc-ing this message to (Randall Holmes and Daniel Dzierzgowski)—because we are interested in the question of the consistency of the constructive version of Quine's NF. The reason why this is an interesting question is that the proof in NF of the axiom of infinity is nonconstructive, and so it might be that constructive NF is much weaker than NF and more easy to prove consistent. Specifically constructive NF proves that V is not Kfinite, but—or so it seems to us—this is not enough for us to interpret Heyting arithmetic.

I have a number of queries. Perhaps some would disappear if i read the article very closely, but email is so temptingly easy! You say on line -5 on page 486 that you make no use of excluded middle in what follows. But it seems to me that your proof of lemma 3 *does* use excluded middle. It certainly *reads* like a case analysis. Am i missing something?

It seems to me that you are claiming that the Kfinite numerals model peano arithmetic. Do i read you right? What worries me about this is that it has seemed entirely obvious to me (and to Randall Holmes and Daniel Dziergoewski) that Kfinite numerals **don't** do this—you need Nfinite cardinals ("adjoin **disjoint** singletons")

Have we been wrong all along?

best wishes

Thomas Forster

`From jbell@julian.uwo.ca Mon Aug 23 14:04:18 1999`

Dear Dr Forster:

It is gratifying to receive a response to one's published efforts: I often feel that they have about as much chance of being read—let alone responded to—as a message sealed in a bottle and cast into the ocean.

Anyway, concerning my paper. The proof of Lemma 3 does indeed argue by cases, but the premises allow this without using excluded middle by supplying the appropriate disjunctions, namely, $y' \in Y \cup \{y\} \iff y' = y \vee y \in Y$, and $x' \in X \cup \{x\} \iff x' = x \vee x \in X$.

Concerning Kuratowski finite numerals. In fact the numerals modelling Peano's axioms in my paper correspond to the *decidable* Kuratowski finite subsets (least family closed under unions with disjoint singletons)as you will see from the definition of "inductive" on the bottom of p. 286. So the claim would be that the "decidable" Kuratowski finite numerals model Peano's axioms. This is further worked out in a sequel to my paper—a copy of which I will send you by steam mail—due to appear in JSL next year.

I know next to nothing about Quine's systems, but it strikes me as odd that the proof of the axiom of infinity in NF is nonconstructive. Is this because the set of natural numbers one gets is automatically well-ordered, so yielding excluded middle?

I take it that your mailing address is the DPMMS—a place I became familiar with during my 30 years of residence in the U.K.

Cordially, John Bell


```
*****************************
John L. Bell
Department of Philosophy
University of Western Ontario
London, Ontario
Canada N6A 3K7
Tel: (519) 661-3453
Fax: (519) 661-3922
Email: jbell@julian.uwo.ca
*****************************
```


`To: ddz@skynet.be, holmes@math.idbsu.edu, jbell@julian.uwo.ca, tf@dpmms.cam.ac.uk`

Thanks for prompt reply. (I'm cc-ing this to the lads too) Your "decidable Kuratowski-finite" sets are those that Daniel Randall and i call N-finite. Now i believe you, and i can see why your case analysis is ok. In fact i proved this result too, and so did Daniel Dziergowski independently of either of us. Do you know any of his work? Naturally what we (and i imagine you too) would like to be sure of is that the disjointness condition really is necessary. If it is—and i suspect that that is the burden of Mawanda - Chisala—then it looks very likely that constructive NF does not interpret Heyting arithmetic (as i suspect!) Can you say anything about that

v best wishes

Thomas Forster

Dear All,

I'm wondering whether in the absence of an Nfinite set whose double complement is the universe one really can prove that for every Nfinite set there is something which is not in it; I'm thinking that one might be able to prove that for any Nfinite set A there is an Nfinite cardinal which is not in A (under the assumption that there is no Nfinite set whose double complement is V).

–Randall Holmes

`From t.forster@dpmms.cam.ac.uk Sat Feb 05 10:58:00 2000`

Andy,

i was very struck by the hints you were throwing out about how much could be done with Kfinite sets—as opposed to what Daniel Dz calls 'Nfinite' sets (which are Kfinite and discrete—addition of disjoint singletons)

It occurs to me that if one is to DO anything with these rectypes one needs a nontriviality condition of some kind. Classically this is of course "every natural number has a successor". What i'm not clear about is what the nontriviality condition is for Kfinite sets. "Every Kfinite set is disjoint from a singleton"? But then the same notriviality condition holds for Nfinite sets too, and we may as well use them instead.

I'm cc-ing this to PTJ too, in the hope that between you you will be able to say somthing enlightening to me! Thanks

Thomas

`From p.t.johnstone@dpmms.cam.ac.uk Sat Feb 05 12:31:08 2000`

Dear Thomas,

For me, the whole point of K-finiteness is that it makes sense without assuming any axiom of infinity: because it's a local definition (i.e. determines whether a set is finite by looking inside its own power-set), you don't need to assume the existence of a "set of all finite sets". Indeed, it's actually unreasonable to expect the "set of all K-finite sets" to exist, in the same sense that N is the set of all N-finite sets.

If you want to know how to count with K-finites, without assuming any axiom of infinity, then you should read Peter Freyd's unpublished paper on

"Numerals and Ordinals", or (probably easier) my Elephantine version of it. Ask me if you want a copy of this.

Peter

`From t.forster@dpmms.cam.ac.uk Sat Feb 05 15:35:54 2000`

I would like a copy of the elephantine version, if you think i'd find that helpful. I think i can put more clearly what my concerns are. The nontriviality condition i spoke of is really nothing more or less than an axiom of infinity: that there should be enuff of these damned things around for everything to make sense. I think i need to know what the appropriate version of this is for Kfinite sets, so that i can tell whether or not constructive NF appears to prove it.

I shall come looking for the elephant soon.

Thanks

Thomas

I've picked Peter's brains, and i think i now understand what needs to be done. Peter says that the nontriviality condition is that disjoint unions should be defined. And presumably cartesian products as well. In the NF context what this means is that there should be a type-level pairing function **defined on Kfinite sets**. It seems to me unlikely in the extreme that the failure of V to be Kfinite allows one to define such a pairing function, since if there really is such a function then the inductively defined set containing the empty set and containing the pair $\langle x, x\rangle$ whenever it contains x is presumably an implementation of arithmetic. But i'll check it.

Conversation with Jeff Egger. He says that Freyd has the term 'Russell Finite' for the following idea:

Given X, consider the *finite stages* of X. First stage is the empty set, hit a stage S by adding to it all sets of the form $y \cup \{z\}$ for $y \in S$ and $z \in X$. The inductively defined set containing emptyset and closed under this operation may or may not contain a fixed point. If it does, we say X is russelfinite. Subsets of rfinite stes are rfinite. Not quite the same as being subfinite sez jeff.

Richard Squire

`From t.forster@dpmms.cam.ac.uk Sun Nov 05 16:23:58 2000Subject: S-B`

I've been amusing myself going over my file *i*NF.tex. Our conversation the other day about S-B is germane to this. It reminds us that constructively $\leq$ between cardinals of Kfinite sets is not provably antisymmetrical!

Is Nfinite the largest subset of Kfinite obeying this antisymmetry?

Thomas

From butz@vip114.it-c.dk Tue Jul 02 10:19:13 2002

To: Thomas Forster ¡T.Forster@dpmms.cam.ac.uk¿

Subject: intuitionistic NF

Thomas,

I started thinking a little about intuitionistic NF. It occured to me that one can use some old results of Pitts (proved in his wild and young days as a topos theorist) to reduce it to a hopefully much simpler (and feasible) problem.

Consider consistent theories T_1 T_2 (in appropriate signatures), and suppose that both T_2 is a conservative extension of T_1, i.e., for sentences in the smaller T_1-language both T_1 and T_2 prove exactly the same sentences. Then for *any* (consistent) extension T of T_1, the theory $T \cup T_2$ is still conservative over T.

(The proof is an almost trivial argument using the fact that Pitts' Phi functor sends conservative maps of Heyting categories to open surjections of toposes, and the latter are stable under pullbacks.)

The result is obviously wrong classically.

This suggests that it is enough to prove the following (either classically or intuitionistically): T_1 is the empty theory in the language of countably many sorts and binary relation symbols $\in$ between 'successive' sorts. T_2 is the theory in the signature above extended by function symbols relating successive sorts and saying that those functions are $\in$ isomorphisms. Let us denote these theories better by E (for empty) and TSA (for type shifting ($\in$) automorphism). If TSA is conservative over E then type theory (extensionality plus comprehension) union TSA is conservative over type theory, hence consistent and Specker's result applies (the argument seems not to be related to the use of classical logic).

What do you think? The latter sounds indeed feasible.

Carsten

From Sergei Tupailo

Dear Professor Forster,

> (We agreed, while Boffa was still alive, that it should be called *i*NF)

I'm fine with the acronym *i*NF, and I actually called it so in my talks. To that end, I got a remark from one of the listeners (Harvey Friedman, Columbus OH) that "*i*NF" is a bad choice, since it reminds of "infimum" and "infinity". Then I was advised to follow the standard practice of denoting the classical and intuitionistic versions of the same theory T by T^c and T^i, respectively, where the superscript can be omitted if it's understood by convention. So my "NFi" was an ASCII abbreviation for NF^i. .

> > If one wants to prove |NFi| = |NF|, of course one thinks about the
> > double-negation translation – what other methods are there for this
> > strength? The double negation translation, applied to NF directly, fails only on the
> > Extensionality axiom – that's a serious problem, and that's easy to
> > see.

> I think you also have a more immediate problem arranging for the
> universe to be a set. Why should the collection of hereditary stable sets
> exist and be a stable set?

Why should we choose the universe to be anything hereditary? But it really depends on the details of how one tries to do things. Perhaps you wouldn't object that the double negation translation works for SF—that's an easy fact, pointed, for example, by Marcel Crabbé in one of his papers.

> > Recently I came across Grishin's and Boffa's 1973 result that NF is
> > equiconsistent with NFU+"O is Cantorian", where O is the set of empty
> > sets (see Boffa's 77 JSL paper). Luckily, the statement "O is Cantorian"
> > seems to survive the double negation translation, i.e. it looks like I can prove its
> > double negation translation in NFi, – this requires some work, but this seems to be
> > true. However, even if true, as it stands this would give only
> > interpretation of classical SF+"O is Cantorian" in NFi, where SF is the
> > part of NF without Extensionality at all. The question about the
> > strength of SF+"O is Cantorian" has to be investigated further, I don't
> > know the answer.

> In principle you are of course quite right: one could seek a double
> negative interpretation of a system classically equivalent to NF, such as
> the one you consider. My guess is that you will find that the
> equivalences between the two theories will rely too much on classical
> logic for the strategy to succeed: i sense that you are aware of this
> danger!

Let's talk in general. Assume we have two first-order recursively axiomatizable classical theories, T_1 and T_2, and T_3 be simply the intuitionistic version of T_2. Assume:

(1) it has been proved that $Consis(T_1) \longleftrightarrow Consis(T_2)$;

(2) the (pure) double negation translation works for T_2, i.e. there is an embedding of T_2 into T_3 using this method.

(2) implies that Consis(T_2) $\longleftrightarrow$ Consis(T_3), this fact being provable in HA, Heyting Arithmetic, (PRA, primitive-recursive arithmetic, would suffice). So,

(3) we have a proof that $Consis(T_1) \longleftrightarrow Consis(T_3)$.

The worst thing which could happen here is that, although $Consis(T_1) \longleftrightarrow Consis(T_2)$ is an arithmetical statement, our proof in (1) could have been done in a theory T much stronger than PA or HA, and maybe classical. So our result (3) could have been established in T, which might not be what we wanted. However, usually this doesn't happen: usually the relevant mathematical results translate into $Consis(T_1) \longleftrightarrow Consis(T_2)$ being provable in PRA, but of course each particular case requires its own examination.

> As for your question about wellordering, you have
> to be very careful, because there are various constructively inequivalent
> notions corresponding to the classical cooncept.

Of course.

> What i might do, if you (and
> Gregori M to whom you copied your message and to whom i'm going to copy
> this) is interested, is the following. Some years ago i discussed with

> Andres Blass the possibility of writing a survey/background article on
> constructive NF for the Bulletin of Symbolic Logic. Blass is no longer an
> editor, but i might write up my notes on this and send it to the BSL
> anyway. If you (and Gregori) would be interested in seeing draughts of
> this document i would be delighted to show it you - if you promise some
> useful feedback!

I can promise the amount of feedback I am able to give.

Very best wishes,

Sergei

On Sat, 30 Apr 2005, Sergei Tupailo wrote:

> Dear Professor Forster,
> > and i think i had somehow got the wrong impression, as it didn't sound as
> > if what you were doing was particularly constructive.
>
> Not in this project, as it was thought of originally. However, I keep
> your question about NFi (intuitionistic NF) in mind. Surely, the known
> proof of the Infinity axiom in NF does seem to use classical logic
> essentially, but this fact alone is not sufficient to expect that NFi is
> weaker than NF. Conversely, problems one encounters if trying to build a
> model of NF seem to be independent of whether the logic is classical or
> intuitionistic, I don't see any reasons why to build a model of NFi
> could be any easier than to build a model of NF.

tf writes

There are several reasons. One is that there is a possibility of a realizability model of *i*NF (We agreed, while Boffa was still alive, that it should be called *i*NF). This is because there is an obvious lambda term corresponding to raising the type of a formula. Another reason is the very classical nature of the proof of the axiom of infinity. I see no way of proving in *i*NF that there is an implementation of Heyting Arithmetic. This suggests that *i*NF is much weaker.

> If one wants to prove |NFi| = |NF|, of course one thinks about the
> double-negation translation–what other methods are there for this
> strength? The double negation translation, applied to NF directly, fails only
on the
> Extensionality axiom – that's a serious problem, and that's easy to
> see.

tf writes

I think you also have a more immediate problem arranging for the universe to be a set. Why should the collection of hereditary stable sets exist and be a stable set?

Sergei writes:

> " A hope to bypass this problem could be to apply the double negation translation to some other, more double negation friendly, system, which (using classical methods !) have been (or could be) shown to be equiconsistent with NF. To explain what I have in mind, here is an example:
>
> Recently I came across Grishin's and Boffa's 1973 result that NF is equiconsistent with NFU+"O is Cantorian", where O is the set of empty sets (see Boffa's 77 JSL paper). Luckily, the statement "O is Cantorian" seems to survive the double negation translation, i.e. it looks like I can prove its double negation translation in NFi, – this requires some work, but this seems to be true. However, even if true, as it stands this would give only interpretation of classical SF+"O is Cantorian" in NFi, where SF is the part of NF without Extensionality at all. The question about the strength of SF+"O is Cantorian" has to be investigated further, I don't know the answer."

tf writes

In principle you are of course quite right: one could seek a double negative interpretation of a system classically equivalent to NF, such as the one you consider. My guess is that you will find that the equivalences between the two theories will rely too much on classical logic for the strategy to succeed: i sense that you are aware of this danger!

Sergei writes:

> "Related question: Does intuitionistic NF prove Infinity? If not Infinity, is there anything similar it's known to prove? Does it prove "V cannot be well-ordered"? If NFi is able to prove at least something somehow related to Infinity, this again could give rise to situations as described above."

tf writes

Well it all depends on what you mean by infinity. *i*NF certainly proves that not every set is Kuratowski-finite, but that's not the same as proving that there is a genuinely inductively infinite set: that inference uses classical logic. As for your question about wellordering, you have to be very careful, because there are various constructively inequivalent notions corresponding to the classical cooncept.

Sergei writes:

> "P.S. Yesterday, when pondering about these issues, I seem to have proved that NFi also refutes a certain version of the Axiom of Choice. This seemed like another argument to expect that NFi has a pretty big strength. Then I looked into your article "Quine's NF, 60 years on", downloaded from your webpage http://www.dpmms.cam.ac.uk/ tf/

, and there on p.6 2nd paragraph seems to be something like a confirmation of this. Is it known that NFi proves $\neg\exists n \in \mathbb{N} V \in n$? Something like this might be enough to claim that NFi is pretty strong. A comment I ought to make to that place is that (the double-negation translation technique tells us that) in intuitionistic logic one can achieve the same strength by using only negative axioms (which have no existential quantifiers at all). Therefore, for the strength it's not necessary to have "there exists an infinite set", something like "not-not there exists an infinite set", i.e. "not every set is finite" (NB: it might be not exactly this, one has to see the details in order to make a clean argument) might be enough. Can you tell me exactly what that result (you're mentioning on p.6 l.13-14) is, or (even better) can you give me a reference or a file? When applying not-not's in a careful way, this might lead to a proof that NFi has the strength at least of classical SF + "V is infinite", which, I hope, has the strength of Simple Type Theory with Infinity."

tf writes

I'll be happy to show you a proof of this. What i might do, if you (and Gregori M to whom you copied your message and to whom i'm going to copy this) is interested, is the following. Some years ago i discussed with Andreas Blass the possibility of writing a survey/background article on constructive NF for the Bulletin of Symbolic Logic. Blass is no longer an editor, but i might write up my notes on this and send it to the BSL anyway. If you (and Gregori) would be interested in seeing draughts of this document i would be delighted to show it you - if you promise some useful feedback!

very best wishes

I'm copying this to Randall Holmes and Marcel Crabbé who i think will be interested too

(End of correspondence with Sergei)

What if there is a dense Nfinite set, X, say. $(\forall y)(\neg\neg(y \in X)$? This does not imply $(\forall y)(\exists x \in X)(\neg\neg(y = x))$. But in the case of interest to us X is Nfinite. Can we prove by induction on Nfinite sets that $(\forall y)(\neg\neg(y \in X))$? No, because $\neg\neg(p \vee q)$ does not imply $\neg\neg p \vee \neg\neg q$. This draws our attention the fact that there may be more stuff in $\sim\sim\{x, y\}$ than in $\sim\sim\{x\} \cup \sim\sim\{y\}$. (there might be things that will always eventually turn out to be to x-or-y but will not always turn out to be x nor will they always turn out to be y.

We need a stronger notion of denseness.

Let us say that X is *strongly dense* if $(\forall y)(\exists x \in X)(\neg\neg(y = x)$. Let X and Y be two discrete strongly dense sets are the same size. Then

$$\lambda x \in X. \bigcup (Y \cap \{x' : x' \sim x\})$$

is a bijection between them. My guess is that if X is a strongly dense discrete subset of V then $\mathcal{P}(X)$ is a strongly dense discrete subset of $\mathcal{P}(V)$. But $V = \mathcal{P}(V)$ so this should give rise to a model of classical NF.

$\{\{x : p = \{y\}\} : p \subset \{y\}\}$

From Alex Simpson

In intuitionistic set theory (the exact variant doesn't much matter), many classically equivalent descriptions of the set of real numbers give rise to different notions of intuitionistic reals. For example, there are different definitions of "Cauchy reals", obtained by varying the notion of Cauchy sequence (of rationals) and perhaps also the notion of equivalence between Cauchy sequences. Nevertheless, there is a widely accepted 'correct' intuitionistic definition, according to which the rate of convergence of a Cauchy sequence must be given by a function (from natural numbers to rationals), in which case the definition of the equivalence of Cauchy sequences is uncontroversial (the two most natural alternatives agree). In fact, an equivalent and often more convenient approach is to assume a fixed rate of convergence (e.g. $1/2^i$). Thus one can define the set of Cauchy reals to be the set of equivalence classes of such fixed-rate-convergent Cauchy sequences of rationals.

I have some questions concerning such Cauchy reals, and other related notions of real number.

1. In Troelstra and van Dalen's "Constructivism in Mathematics", the "Cauchy completeness" of the Cauchy reals is proved by defining a "Cauchy sequence of reals" to be given by a sequence of representative Cauchy sequences of rationals. However, a more natural definition of Cauchy sequence of reals is to take instead sequences of reals themselves (i.e. sequences of equivalence classes of Cauchy sequences of rationals). Without number-number choice (by which I mean the, classically provable, Axiom of Choice for $\forall\exists$ prefixes that quantify over natural numbers—often called AC_{00}) the more natural notion of sequence is apparently more general than version using representatives. In fact, if the more natural definition is used, it does not seem to be possible to prove that the Cauchy reals are Cauchy complete. My first question is: does anybody know a model for some reasonable intuitionistic set theory (e.g. a topos) in which the Cauchy reals are *not* Cauchy complete in this sense?

2. In an impredicative set theory, one can also define a natural notion of Dedekind real (again there is one 'correct' definition - namely R_d in Troelstra and van Dalen). The set of Dedekind reals *is* Cauchy complete. Thus one can also define the "Cauchy-completed reals" R_cc as the intersection of all Cauchy-complete subsets of R_d containing the rationals Q. Easily, the Cauchy reals, R_c, embed in R_cc. Thus one has injections:

$$R_c \to R_{cc} \to R_d$$

Given number-number choice, both inclusions are isomorphisms. I know models (e.g. sheaves on $\Re$) in which the inclusion $R_cc \to R_d$ is proper (i.e. it is not an isomorphism). A reformulation of Question 1 is whether there exists a model in which the other inclusion, $R_c \to R_{cc}$, is proper. Question 2 is: has anyone seen the Cauchy-completed reals (or something equivalent to them) defined before? Any references would be very welcome.

3. There is an alternative take on the inclusions in 2. One can define R_d as: convergent round filters of proper rational intervals (a proper interval is a pair (q_1, q_2) with $q_1 < q_2$; a round filter is a filter w.r.t. strict (i.e. proper) inclusion of intervals; convergent simply means that for any epsilon there is an interval of width ¡ epsilon in the filter). One can also exhibit R_c explicitly as a subset of R_d as the set of all "countably-based" such filters (where countably-based means that there exists a function N $\to$ the filter giving a filter base). Question 3 is: does there exist similar explicit description of R_{cc} as those filters in R_d satisfying some good property?

4. The above questions are motivated by a geometrically-based approach to axiomatizing the real numbers that Martin Escardo and I have been working on. When interpreted in intuitionistic set theory, our axioms yield the Cauchy-completed reals. The above roundabout construction of the Cauchy-completed reals via Dedekind reals makes crucial use of impredicative notions such as powerset and intersection of all subsets. Is this essential? More generally, in predicative intuitionistic set theories, like Aczel's CZF, is it possible to define *any* reasonable Cauchy-complete notion of real number *without* first assuming number-number choice?

```
--
Alex Simpson, LFCS, Division of Informatics, University of Edinburgh
Email: Alex.Simpson@dcs.ed.ac.uk          Tel: +44 (0)131 650 5113
FTP: ftp.dcs.ed.ac.uk/pub/als             Fax: +44 (0)131 667 7209
URL: http://www.dcs.ed.ac.uk/home/als
```

From Carsten Butz

Dear Thomas

cc: Randall

Administrative duties have kept me away from anything serious for way too long. Last Thursday and Friday, however, I had a very pleasant seminar with four guests, among others Martin Hyland, about realizability and Dialectica interpretations, which reminded me of the fun research is all about.

I tried to look at the stuff of the Russian guy, who recently (read: about 3 years ago) claimed equiconsistency of NF and *i*NF, but as far as I am tell, this is not correct (read: I don't really understand what he is doing).

As to the "ideas": I don't remember Pitts, but I will have a look at my notes to see, whether there was something in that direction. There are, however, two more "promising" ideas, which I haven't really pursued. The one is yours about realizability models for type theory, where one would have a realizer for a type-shifting morphisms almost automatically. The other one is topos theory, where one builds a model of a certain "simple" type theory in a certain topos, and unwinding the construction yields a model of the type theory one really is interested in. You should compare this to a group in the topos of simplicial sets (for example), which, once unwound, is the same data as a simplicial group, a reasonably complex structure with all sorts of morphisms and identities relating the many groups that make up the simplicial group. The right (?) topos to look at could be something like (pre-) sheaves over the groupoid consisting of countably isomorphic copies of one object. However, when unwinding the notion of an internal model of a one-sorted theory with an epsilon relation, I haven't really though too much about, what kind of properties the model inside the topos should have, so that its externalization is a model of type theory with an type shifting automorphism. The worst case scenario, not very unlikely, is, that you need an *internal* model of NF, to get an *external* model of type theory (now, while writing these lines, this sound plausible). However, there is also the possiblity to construct something weaker, say a sequence of models where at each successor node you have the regular power-set, and then do a countable sequence of cardinal collapses, and see, what one gets in the limit. Again, I am not sure whether this type of arguments (working in toposes boils down to forcing arguments) really helps.

The last thing I probably mentioned goes in a slightly different direction: In Sets there are no complete (or co-complete) categories except the trivial ones: complete posets. This changes if you are in an intuitionistic universe (though one has to be careful about, what completeness really means). Realizability models contain such gadgets, and Martin Hyland knows more about this than I do. Such categories can probide non-trivial models of polymorphic (not the word you want to use) type theories, thus, this is also an area to look for models of intuitionistic NF. However, I haven't looked at this. Andy knows a lot about these things, and maybe this was where I mentioned him.

For the moment that is all I have to say. Thanks for the invitation. If I remember correctly, I also have a "standing invitation" from Andy, but I had no time to actually visit Cambridge. One of our phd students, Bodil Biering, was in Cambridge earlier this year for her half year visit abroad, you must have met her. Even then I didn't find time to go to Cambridge to find out how she was doing, really bad.

All the best
Carsten

tf to Carsten Butz

Carsten,

I am sorry i have been out of touch for so long. I have very happy memories of our time in Copenhagen together—which says something because it was so hot i thought i was going to die. In fact it's partly because of European summers that i am at present in New Zealand → estivating! (there's a good obscure word for you). I get back at the end of october. The reason i am pestering you now is that Randall and I are planning to write a survey article for the Bulletin of Symbolic Logic covering everything that is known about constructive NF. Have you had any more thoughts? I still have a folder of correspondence with you on this subject. You may remember that it started because you had the idea that an article of Pitts could be useful in this context. I don't seem to have a reference for this article. Can you put your hand on it easily? Then we started thinking about constructive tangled type theories. I think i can probably persuade Randall to sort that out (He invented tangled types after all!) Anyway, i would be grateful for any hints you feel able to give us. My research group here in Cambridge has some funding, so if you find the time to come over here to talk i should be able to find some money to support you.

How are you anyway?

A Hiatus Here

This last seems most unlikely, since if *i*NF had the existence property (as conjectured), then by the existence property we would have a term t_x and a proof that $\mathrm{Nfin}(x) \to t_x \notin x$. (As before, notice that because of the equivalence with type theory, this t_x must be one type lower than x (otherwise this would not be a theorem of the underlying intuitionistic type theory!)) Then the least set containing $\emptyset$ and closed under $\lambda x.(x \cup \{t_x\})$ would give an implementation of arithmetic, taking that set to be $\mathbb{N}$ and the function $\lambda x.(x \cup \{t_x\})$ to be Successor.

Implement
$0 =: \emptyset$;
$S(x) = x \cup \{t_x\}$;
$\mathbb{N} =: \bigcap\{x : 0 \in x \wedge S``x \subseteq x\}$.[10]
The tricky part is always to show that S is one-to-one. Define $<$ as an inductively defined set—$x < y$ if $y = S`x \vee S`x < y$—so that

$$< \; = \; \bigcap\{R : S \subseteq R \wedge (\forall uv)(\langle u, v\rangle \in R \to \langle u, S(v)\rangle \in R)\}$$

Prove by induction on x that $(\forall y)(x < y \vee x = y \vee y < x)$. This is certainly true if $x = 0$; Suppose $(\forall y)(x < y \vee x = y \vee y < x)$. We wish to infer the same for $S(x)$. Think of an arbitrary y. By induction hypothesis we have $x < y \vee x = y \vee y < x$. In the last two cases we infer $y < S`x$. In the first case, $x < y$ is $y = S`x \vee S`x < y$. and these are the two missing cases in the conclusion.

Now suppose $x \cup \{t_x\} = y \cup \{t_y\}$. In any case we have $x < y \vee x = y \vee x > y$. Suppose $x < y$. Then $x \subseteq y \to t_y \notin x$. But $t_y \in x \cup \{t_x\}$ so $\neg\neg(t_y = t_x)$, so $\neg\neg(t_y \in y)$. But we know that $t_y \notin y$. The case $y < x$ is excluded similarly. So $x = y$.

[10] Note the similarity to the Von Neumann implementation of naturals: if we have foundation then t_x can be taken to be x itself.

This tells us that S is 1-1. Trichotomy also tells us that $=$ between members of $\mathbb{N}$ is determinate: $(\forall n \in \mathbb{N})(\forall xy \in n)(x = y \vee x \neq y)$.

(Notice that the converse is easy, at least classically: If we have an implementation of arithmetic in NF, then take t_x to be the first member of $\mathbb{N} \setminus x$.)

If we assume something slightly stronger than that there is an implementation of Heyting Arithmetic into *i*NF, namely that the cardinals of Kfinite sets form such a model, then we deduce excluded middle as follows. In Heyting arithmetic we have $n = m \vee n \neq m$, which implies here that the set of Kfinite cardinals is discrete. If we want to interpret natural numbers as cardinals of Kfinite sets this corresponds to it being determinate whether or not two Kfinite sets are the same size. In particular for any old x and y we must have $|\{x, y\}| = |\{x\}| \vee |\{x, y\}| \neq |\{x\}|$ which clearly will imply $x = y \vee x \neq y$. (I think this is the point of Mawanda and Chisala) This makes this a rather unnatural version of the axiom of infinity for *i*NF. This is probably why Dzierzgowski has for some time believed that the correct version of the axiom of infinity for *i*NF is the assertion that the cardinals of Nfinite sets form a model of Heyting arithmetic. This is a strong assumption all right, but it doesn't seem to have any strong consequences for the logic.

There is a proof in NF that if everything is a term then there is no choice function on the set of all pairs. Think about reproducing this proof in *i*NF if *i*NF has the existence property. The argument in the classical case relied on the transposition (t_1, t_2) where t_1 and t_2 are two closed terms. Intuitionistically this permutation is defined only when t_1 and t_2 are—to coin a phrase—*isolated*: $(\forall x)(x = t_1 \vee x \neq t_1)$. Are there *any* such terms? $\emptyset$ is almost like this—anything $\neg\neg$ equal to it is equal to it, but that is weaker... The moral seems to be that there is not much mileage to be made out of this idea. After all, there is theorem 14 that says that there are no isolated sets unless the logic is classical.

```
From ablass@math.lsa.umich.edu Sat May 25 10:06:24 1996
Received: by emu.pmms.cam.ac.uk (UK-Smail 3.1.25.1/1); Sat, 25 May 96 10:06 BST
Received: from zubrovka.math.lsa.umich.edu by math.lsa.umich.edu (8.7.1/2.3)
        with ESMTP id FAA17532; Sat, 25 May 1996 05:09:14 -0400 (EDT)
Date: Sat, 25 May 1996 05:09:13 -0400 (EDT)
From: Andreas Blass <ablass@math.lsa.umich.edu>
To: Thomas Forster <T.Forster@pmms.cam.ac.uk>
Subject: Re: Kfinite sets
In-Reply-To: <m0uKnp3-0003WdC@emu.pmms.cam.ac.uk>
Message-ID: <Pine.SUN.3.91.960525050724.6418B-100000@zubrovka.math.lsa.umich.edu>
MIME-Version: 1.0
Content-Type: TEXT/PLAIN; charset=US-ASCII
Status: RO
```

Dear Thomas,

I believe your conjecture that hereditarily K-finite implies N-finite can be proved as follows.

First, I claim that equality between hereditarily K-finite sets is decidable, i.e., either $x = y$ or not $x = y$. This is proved by induction on hereditarily

K-finite sets x (for all y simultaneously) as follows. (I assume that the definition of "hereditarily K-finite" is something like "the smallest class containing all K-finite subsets of itself", so that such inductions are justified.) Given (hereditarily K-finite) x and y, we have, for all members x' of x andmy' of y, that $x' = y'$ is decidable, by induction hypothesis. But decidability is preserved by quantification over K-finite sets and by conjunction, so we also have decidability of

$$(\forall x' \in x)(\exists y' \in y)x' = y'$$

and

$$(\forall y' \in y)(\exists x' \in x)x' = y'$$

. That is, we have decidability of $x = y$.

To finish the proof, I claim that K-finiteness of a set z plus decidability of equality between its members implies N-finiteness of z. (This is undoubtedly well-known, but I'll give the proof anyway for completeness.) Proceed by induction on K-finite sets, the case of the empty set being trivial. So suppose $a \cup \{x\}$ has decidable equality between all its members (where a is a K-finite set for which the result is known to hold). In particular, each member of a is either equal to x or not. Using again that quantification over K-finite sets preserves decidability, we find that x is either in a or not. So $a \cup \{x\}$ is either just a (which is N-finite by induction hypothesis, because equality between its members is decidable) or the disjoint union of a and $\{x\}$, which is N-finite by definition of N-finiteness. That completes the proof.

Though it's not relevant to this argument, I might mention that the converse of the last paragraph works also: Every N-finite set is K-finite and equality between its members is decidable. The proof is a routine induction on N-finite sets.

Best regards,
Andreas

Build a tree below $|V|$ by putting below each node β a cardinal $|\iota``x|$ iff $|\mathcal{P}(x)| = \beta$ **as long as x is Kfinite**.

Let us hope that we can show that

$$(\forall x, y)((\text{kfin}(x) \wedge \text{kfin}(y) \wedge \text{kfin}(\mathcal{P}(x)) \wedge |\mathcal{P}(x)| = |\mathcal{P}(y)|) \to |x| = |y|).$$

Since by Cantor's theorem $|\iota``x| \not\geq_* |\mathcal{P}_{kfin}(x)|$ (so in particular $|\iota``x| \neq |\mathcal{P}_{kfin}(x)|$) the successive initial segments of the tree are N-finite. The whole tree cannot be infinite (why?) so we should be in with a chance of finding a bottom element.

So suppose there is n such that $S(n) = \emptyset$. Then there can be no x such that $(\forall y \in x)(x \setminus \{y\} \in n)$. The end of this trail will be that for any old $x \in n$ we have $(\forall y)(\neg\neg(y \in x))$

One trick that may be useful is this. Try: 0 is implemented as $\{\emptyset\}$, and

$$S(n) =: \{y : (\exists x \in n)(\exists z)(y = x \cup \{z\})\}$$

Notice that these integers are $\subseteq$-cumulative. Is there a last one? Notice that there is nothing to stop us forming the union of all those integers that are not equal to FIN and reasoning about the difference between that and FIN. Apparently this implementation of naturals is called *Church* naturals.

To get sensible Frege-implementations of $\mathbb{N}$ you have to restrict attention to Kfinite sets x such that $(\forall y, y' \in x)(y = y' \vee y \neq y')$. This is because when you delete y from x you also delete everything $\neg\neg$ equal to y. Why does this matter? The point is, it stuffs up *both* definitions of Frege successor in *i*NF. We set $S`n$ to be $\{x : (\exists y \in x)(x \setminus \{y\} \in n)\}$ or $\{x : (\forall y \in x)(x \setminus \{y\} \in n)\}$. To keep things simple, consider the case $n = 0$, and consider an inhabited set X with lots of members all $\neg\neg$ equal to each other. According to *either* of these definitions X should belong to 1, which it doesn't.

PTJ says: $\mathcal{P}(x)$ can be Kfinite even if x isn't: think of a p st $\neg\neg p$ and $p \to$ everything is boolean. Then $\{x : x = a \wedge p\}$ is not Kfinite but its power set is.

Dense N-finite sets again

I think we can prove by induction that every Nfinite set is even (has a partition into pairs) or odd (has a partition into pairs + one singleton) but not both. Let A and B be dense Nfinite with A odd. Then $\neg\neg(B$ is odd). But B is odd or even so it must be odd. Even similarly.

Can we show that any two Nfinite sets that are notnotequal are in bijection? It would be an induction. Suppose A, B, $A \cup \{a\}$ and $B \cup \{b\}$ are all Nfinite with $a \notin A$, $b \notin B$ and $\neg\neg(A \cup \{a\} = B \cup \{b\})$.

Working inside $\neg\neg \ldots$ We get $\neg\neg(A \setminus \{b\} = B \setminus \{a\})$. and thence a bijection between $A \setminus \{b\}$ and $B \setminus \{a\}$ and finally a bijection between $A \cup \{a\}$ and $B \cup \{b\}$.

(We are going to need that if $A \cup \{a\}$ is Nfinite then $A = (A \cup \{a\}) \setminus \{a\}$.)

So we are working towards a situation wherein all dense Nfinite sets are notnotequal and all the same size. So ask next: "what can we say about sets that are in bijection with a dense Nfinite set?" Sse A is dense Nfinite, and f is a bijection. Is $f``A$ dense? We trade on the fact that $x = y \longleftrightarrow f(x) = f(y)$ and, for all $x$, $\neg(\forall y \in A)(x \neq y)$ to obtain for all $x$, $\neg(\forall y \in A)(f(x) \neq f(y))$, which is $\neg(\forall y \in f``A)(f(x) \neq y)$ which appears to say that $f``A$ is dense. But we don't really know what this means, beco's f is defined merely for things that are actually in A. But it does look like something that ought to be true. Of course if f is a permutation we're OK, but there is reason to believe that there aren't any permutations.

Extended Kuratowski-finite

Kuratowski-finite sets. The $\subseteq$-least collection containing the empty set and closed under adjunction: $x, y \mapsto x \cup \{y\}$. PTJ tells me that there is no current name for the weaker, more inclusive, property where one is additionally allowed unions of chains. I am expecting to be able to prove analogously that if Ω has this property then the logic is classical.

For the moment let's call it EKF for 'extended Kuratowski finite'. Pretty clear that a surjective image of an EKF set is EKF. We have to show that $\mathrm{EKF}(\Omega) \to$ TnonD. Does EKF sets obey Linton-Johnstone? I suspect not, co's i've tried and got nowhere. However we might be able to show that Ω is not the union of a chain of Kfinite sets. I think we would need it to be true that if a, b both in the union of a chain of things, then there is a thing that contains both.

Bitwise Representation

...of naturals, using Ω instead of $\{0, 1\}$. How do you describe the *carry*? It might help us in doing CO models constructively.

A filter on Ω?

How about a generic object G s.t. the truth-value of $a \in G$ is precisely a? That would be a kind of ultrafilter. Can we find a set abstract that does that?

$[[a \in G]] = \{x : x = \emptyset \wedge a \in G\}$

That is to say, we seek $G \subseteq \mathcal{P}(\{\emptyset\})$ such that $(\forall a \subseteq \{\emptyset\})(a = \{x : x = \emptyset \wedge a \in G\})$.

Mind you, $\sim\sim\{\{\emptyset\}\}$ is a filter. If $\neg\neg(a = \{\emptyset\})$ and $\neg\neg(b = \{\emptyset\})$ then $\neg\neg((a \cap b) = \{\emptyset\})$ so it's closed under $\cap$. (It's obviously closed under $\supseteq$.)

Do we have a good notion of a quotient? $x \sim_F y$ iff $x \to y$ and $y \to x$ both in F.

$\overline{B}x$ again

Suppose $\mathcal{V}$ is a dense set, $\sim\sim\mathcal{V} = V$, and let a be any set. Then $\overline{B}a = \sim\sim\{x \in \mathcal{V} : a \notin x\}$

If $y \in \overline{B}a$ then $a \notin y$ and $\neg\neg(y \in \mathcal{V})$ so $\neg\neg((a \notin y) \wedge (\neg\neg(y \in \mathcal{V})))$ which is to say $y \in \sim\sim\{x \in \mathcal{V} : a \notin x\}$. And i think the arrows can be reversed.

This means that $\overline{B}``V$ can be injected into the power set of $\mathcal{V}$. Can't do anything with that *at the moment* but put it on one side co's it may yet come in handy.

1-symmetric functions

Write 'D' for double complement. Does D commute with $j(D)$? With $j^2(D)$? $D``(D(x)) = D(D``x)$?? That would help...

LHS:

$y \in D``(D(x))$ iff

$(\exists z \in D(x))(y = D(z))$ iff

$(\exists z)(\neg\neg(z \in x) \wedge y = D(z))$

RHS:

$\neg\neg(y \in D``x)$ iff

$\neg\neg(\exists z \in x)(y = D(z)))$

And these two are the same if x is kfinite but presumably not otherwise. Check it!

And there's always a question of this sort about functions that are 1-symmetric. Classically the only 1-symmetric functions are ... well, i was about to say the identity and complementation, but there's also KV and $K\emptyset$ and possibly even a few others. Should track them down.

19.3.1 pseudoextensional functions

Is the right context to think about pseudoextensional functions such as ι?

20 Miscellaneous Other Proof Theory

Reflect on the following banality. If T is a constructive set theory which admits cut elimination, then T does not refute the existence of V.

In particular, no cut-elimination for Zermelo!! (which we probably knew anyway)

This argument doesn't use the fact that $\mathcal{P}^n X$ can be a subset of x. The point is this: once you've got to the line $\vdash t_x \notin y$, this can only have arise from $\neg$-R, so take it over to the left. Then $t_x \in x \vdash$ doesn't match the output of any sequent rule. (You need cut-elimination and constructivity to be able to reason that any proof must look like that....)

I've looked at constr weaker things that say V isn't a set and the same argument works.

So at the very least, cut-elimination for KF proves con (iNF) - and that's without using the bounding lemma style stuff.

Dear Thomas and Randall,

... Anyway, I don't know if intuitionistic TST + weak ambiguity is consistent. By the way, why are you using the rule $\Gamma \vdash \sigma$ gives $\Gamma \vdash \sigma+$ instead of the axioms $\vdash (\sigma \longleftrightarrow \sigma+)$? I'm not sure they are equivalent (try to prove the axioms from the rule). Notice that the axiom is "symmetric" in sigma and sigma+, while the rule is not. In classical logic, $p \longleftrightarrow q$ is equivalent to $((p \to q) \wedge (\neg p \to \neg q))$. That's why, in classical logic, ambiguity can be axiomatized with non symmetric axioms $\sigma \to \sigma^+ \;\vee\; \sigma^+ \to \sigma$.

Notice that for full ambiguity, you can use the rule $\Gamma \vdash \sigma$ gives $\Gamma \vdash \sigma^*$ or the axioms $\vdash (\sigma \longleftrightarrow \sigma^*)$, where σ^* denotes any variant of σ. The reason is that $(\sigma \longleftrightarrow \sigma)$ is a tautology and $(\sigma \longleftrightarrow \sigma^*)$ is a $(\sigma \longleftrightarrow \sigma)^*$.

In classical logic, Hilbert style, ambiguity can be axiomatized with the rule $\vdash \sigma$ gives $\vdash \sigma^*$, or with the axioms $(\sigma \longleftrightarrow \sigma^*)$. It can also be axiomatized with the usual axioms $(\sigma \longleftrightarrow \sigma^+)$, but I don't know if it can be axiomatized with some "+" rule: $\vdash \sigma$ implies $\vdash \sigma^+$ is of course not correct and I cannot remember the point with $\vdash \sigma^+$ implies $\vdash \sigma$.

Also, I could prove that *in pure predicate calculus* weak ambiguity is not equivalent to full ambiguity, in intuitionistic logic. But I was not able to prove that intuitionistic TST + weak ambiguity is not equivalent to int. TST + full ambiguity. I think it is not but I also think that it is as hard to prove as to prove that it is consistent! More later... Daniel.

and some tho'rts of mine, on wed 5/viii

How do we set up the sequent calculus rules for ambiguity in constructive type theory? Presumably

$$\frac{\Gamma \vdash p}{\Gamma \vdash p^+}$$

Now what about the case where p and p^+ are contraries? This might arise in various ways, through derivations of any of the following

$$\Gamma, p \vdash \neg p^+ \quad \Gamma, \neg p \vdash p^+ \quad \Gamma, p^+ \vdash \neg p \quad \Gamma, \neg p^+ \vdash p$$

Now the case that worries me is (there may be others) is (the following notational variant): $\Gamma, p \vdash \neg p^-$.

This gives $\Gamma, p \vdash \neg p$ by the +-rule. Then
$\Gamma, p, \neg\neg p \vdash$ by ¬-L;
$\Gamma, \neg\neg p \vdash \neg p$ by ¬-R;
$\Gamma, \neg\neg p, \neg\neg p \vdash$ by ¬-L;
$\Gamma, \neg\neg p \vdash$ by contr-L;
$\Gamma \vdash \neg\neg\neg p$ by ¬-R. Then we cut against
$\neg\neg\neg p \vdash \neg p$ to get $\Gamma \vdash \neg p$ which we would expect all along.

Now this occurrence of cut cannot be eliminated. This sounds pretty dire, but—at least in LPC—it isn't. The point is that if we have a constructive proof of $\frac{\Gamma, p}{\neg p}$ we cannot have obtained this by assembling p on the left and leaving $\neg p$ undecomposed on the R. This is beco's we would have had no initial sequent with $\neg p$ on the R. So the $\neg p$ on the R must have come from a p on the left, and given that we can use contr-L.

[day of Jesse and Stephanie's wedding .... Why can we not argue thus
$\Gamma, p \vdash \neg p^-$; then
$\Gamma, p \vdash \neg p$ by the +-rule. Then
$\Gamma, \vdash \neg p, \neg p$ by ¬-R; then
$\Gamma, \vdash \neg p$ by contraction.
What's wrong with that?]

This saves the day for cut-elimination for LPC, but it is clear that the cuts cannot be eliminated from the proof above!

The $\in$ rules for sequent calculus.

These must be

$$\frac{\Gamma \vdash \phi(t)}{\Gamma \vdash t \in \{x : \phi\}} \text{ and } \frac{\Gamma, \phi(t) \vdash p}{\Gamma, t \in \{x : \phi\} \vdash p}.$$

That way we can prove

$$\frac{t \in \{x : \phi\}}{t \notin \{x : \neg\phi\}} \text{ and } \frac{t \notin \{x : \phi\}}{t \in \{x : \neg\phi\}}.$$

... but not—of course!—$\dfrac{t \notin \{x : \neg\phi\}}{t \in \{x : \phi\}}$

The point is that to keep things simple one does not allow oneself to introduce anything like $t \notin \{x : \phi\}$ on either side except by the negation rules. But the restriction doesn't prevent us from proving any of the things we want.

Andre,

I remember years ago you saying that the interpolation theorem was a problem for stratified languages. I have been thinking about this recently in connection with cut-elimination for type theory with ambiguity. (Holmes is here and we are going to prove the consistency of constructive NF!) Consider the following situation. We have a proof that $\phi \to \phi^+$, and therefore of $\phi \to \phi^n$ for n as large as one wants, in particular for n so large that ϕ and ϕ^n have no predicate letters in common (all the $\in_k$ are different predicates of course!). If we apply interpolation to this we get an absurd result. Presumably this shows that interpolation fails. But how can this be? Interpolation follows from cut-elimination, and didn't someone (Takeuti?) prove cut elimination for simple type theory (or perhaps it was for type theory without the $\in$-rules)

Have you thought about this

very best wishes

Thomas

21 A factoid

There is a standard trick in classical propositional logic for taking a theory T and a theorem A of T and axiomatising T in such a way that A is an independent axiom of the new axiomatisation. The idea is that one can then consistently replace A by $\neg A$ and maybe something interesting will happen.

I am going to need a constructive version of this result, so i'll spell it out.

Let T and A be as above. We also assume (for the avoidance of triviality) that $\neg\neg A$ is not a thesis of constructive first-order logic.

We axiomatise T with the axiom A and the scheme which we write '$\Sigma_B A \to B$', of conditionals $A \to B$, where the Bs are the theorems of T. Suppose A is not independent, and that it follows from the scheme $\Sigma_B A \to B$. Then, for some finite set of Bs, we have

$$\bigwedge_B (A \to B) \vdash A$$

where the turnstile means provability in constructive first-order logic. But $\bigwedge_B(A \to B)$ is equivalent to $A \to \bigwedge_B B$ whence

$$(A \to \bigwedge_B B) \vdash A.$$

and

$$\vdash (A \to \bigwedge_B B) \to A.$$

Now

$$\vdash ((A \to \bigwedge_B B) \to A) \to \neg\neg A$$

so so (by *modus ponens*)

$$\vdash \neg\neg A,$$

contradicting assumption. So A is an independent axiom.

■

The idea is to do this in a constructive setting where we replace A by $\neg A$ as allowed, but retain as an axiom something *weaker* than A. To this end we say that a formula in $\mathcal{L}(T)$ that is classically equivalent to A but does not imply $\neg\neg A$ constructively is a *weakening* of A, and we will denote such a formula by '$w(A)$'. Observe that Glivenko's thereom that in propositional logic the double negation of a classical tautology is intuitionistically correct means that in propositional logic any $w(A)$ that is classically equivalent to A must constructively imply $\neg\neg A$, so such weakeners are not to be had in propositional logic; their construction must involve manipulation of the quantifiers. We are obliged to work in first-order logic

Now let us say that a closed formula A is *sufficiently complex* if it admits such a weakening.

In particular if A is sufficiently complex it is stronger than its double negation. Notice that anything of the form $\neg A$ stands a good chance of being sufficiently complex, since it is equivalent to its double negation. Negations of ambiguity axioms are almost guaranteed to be sufficiently complex. But I'm getting ahead of myself.

Let T be a theory, A a sufficiently complex theorem of it. As before we assume (for the avoidance of triviality) that $\neg\neg A$ is not a thesis of constructive first-order logic.

We axiomatise T by A and the scheme $\Sigma_B A \to B$, where the Bs are the theorems of T, as above. So we can replace A in this axiomatisation by $\neg A$. That much is standard. The novel point here is that one can do that while retaining $w(A)$ as an axiom! The proof parallels the foregoing.

Suppose *per contra* that

$$w(A), \Sigma_B A \to B \vdash \neg\neg A.$$

By compactness we need only finitely many of these conditionals on the left, and—as above—we recall that, for any fixed A, any finite conjunction of things of the form $A \to B$ is another thing of that form. Thus we get

$$w(A), A \to \bigwedge B \vdash \neg\neg A$$

for some finite set of Bs whence

$$w(A) \vdash (A \to \bigwedge B) \to \neg\neg A.$$

But now we recall (as above) that, for any X,

$$\vdash ((A \to X) \to \neg\neg A) \to \neg\neg A$$

whence

$$w(A) \vdash \neg\neg A$$

But this contradicts the assumption that A was sufficiently complex.

■

This gives us an operation on theories. Input a consistent theory T, and A a sufficiently complex theorem of T. Output a (constructive) theory T^w which has axioms: $w(A)$, $A \to B$ for all axioms B of T ... **and** $\neg A$. **And** T^w is consistent!

OK, i admit it; this is a trivial factoid. So far it only says that you can "turn" (Sorry, i have been watching too much Le Carré of late) a single theorem. Naturally one would like to be able to turn infinitely many expressions simultaneously. I don't think it will work, tho'.

I don't think i am letting any secrets out of the bag by saying that the medium term project is to tackle a *strongly typed* theory (as it might be T$\mathbb{Z}$T). Suppose T$\mathbb{Z}$T proves some negation of an ambiguity axiom. (I don't suppose for a moment that it does—i have far more faith in Randall than that—but in principle it might). Then you 'turn' all the negations of instances (or disjunctions of such negations) of the ambiguity scheme that T$\mathbb{Z}$T proves. (All such formulæ are sufficiently complex). Then you have a theory that is classically equivalent to T$\mathbb{Z}$T and consistent with the ambiguity scheme, so you add the scheme and then *you can drop the type indices*. This doesn't neccessarily give you *i*NF but it does give you a constructive theory in the language of set theory whose axioms are classically equivalent to it. The interest in the foregoing (if any) lies in the fact that it parallels the *aperçu* that realizability gives us a reason to believe in the consistency of *i*NF that is not at the same time a reason to believe in the consistency of NF. It's a straw in the wind.

This has been at the back of my mind for some time, and is the cause of my firmly-held suspicion that *i*NF is weak.

22 Bibliography

References

[1] Bell, John L. Frege's theorem in a constructive setting. Journal of Symbolic Logic **64** (june 1999) pp 486–8.

[2] B. P. Chisala and M-M Mawanda: Counting measure for Kuratowski finite parts and decidability. Cahiers de topologie et géométrie différentielle catégoriques **32** (1991) pp 345-53. `http://www.numdam.org/article/CTGDC_1991__32_4_345_0.pdf`

[3] Crabbé, M. [1975] Types ambigus. *Comptes Rendus hebdomadaires des séances de l'Académie des Sciences de Paris série A* **280** pp. 1−2.

[4] Crabbé, M. [1976] La prédicativité dans les théories élémentaires. *Logique et Analyse* **74-75-76** pp. 255−66.

[5] Crabbé, M. [1978a] Ramification et prédicativité. *Logique et Analyse* **84** pp. 399−419.

[6] Crabbé, M. [1978b] Ambiguity and stratification. *Fundamenta Mathematicae* **CI** pp. 11−17.

[7] Crabbé, M. [1982a] On the consistency of an impredicative subsystem of Quine's NF. *Journal of Symbolic Logic* **47** pp. 131−6.

[8] Crabbé, M. [1982b] À propos de 2^α. *Cahiers du Centre de Logique* (Louvain-la-neuve) **4** pp. 17−22.

[9] Crabbé, M. [1983] On the reduction of type theory. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* **29** pp. 235−7.

[10] Crabbé, M. [1984] Typical ambiguity and the axiom of choice. *Journal of Symbolic Logic* **49** pp. 1074−8.

[11] Crabbé, M. [1986] Le schéma d'ambiguïté en théorie des types. *Bulletin de la Société Mathématique de Belgique série B* **38** pp. 46−57.

[12] Crabbé, M. Stratification and cut-elimination. *Journal of Symbolic Logic* **56** (1991) pp 213–26

[13] Crabbé, M. On NF U. *Notre Dame Journal of Formal Logic* **33** (1992) pp 112–119

[14] Crabbé, M. The Hauptsatz for stratified comprehension: a semantic proof, Mathematical Logic Quarterly, **40** (1994) pp. 481–489.

[15] Radu Diaconescu. Axiom of Choice and Complementation Proc. AMS **51** (1975) 176–178.

[16] D. Dzierzgowski. Finite sets and Natural Numbers in Intuitionistic TT, Notre Dame Journal of Formal Logic, vol. 37, no. 4 (1996), pp. 585–601.

[17] D. Dzierzgowski. Finite Sets and Natural Numbers in Intuitionistic TT without Extensionality, Studia Logica, **61** no. 3 (November 1998), pp. 417–428.

[18] D. Dzierzgowski. "Le théorème d'ambiguïté et son extension à la logique intuitionniste". Dissertation doctorale. Université Catholique de Louvain, Institut de mathématique pure et appliquée. Janvier 1993.

[19] D. Dzierzgowski. Models of Intuitionistic TT and NF. Journal of Symbolic Logic **60** (1995) pp 640–53.

[20] Forster, T. E. Set Theory with a UNiversal set: Exploring an Untyped Universe. OUP

[21] Thomas Forster "A Tutorial on Constructive NF", in the NF 70th anniversary volume *Cahiers du Centre de Logique* **16** 2009 pp 137–171.

[22] Thomas Forster "AC fails in the natural analogues of L and the cumulative hierarchy that model the stratified fragment of ZF". in *Contemporary Mathematics* **361** 2004.

[23] Glivenko, V. Sur la Logique de M. Brouwer. Bulletin de L'Académie Royal de Belgique, série de la Classe des Sciences **14** (1928) pp. 225–8.

[24] Glivenko, V. Sur quelques points de la Logique de M. Brouwer. Bulletin de L'Académie Royal de Belgique, série de la Classe des Sciences **15** (1929) pp. 183–8.

[25] Holmes, M.R. The equivalence of NF-style set theories with "tangled" type theories; the construction of ω-models of predicative NF (and more). Journal of Symbolic Logic **60** (1995) pp. 178–189.

[26] P. T. Johnstone "Conditions related to De Morgan's law" in "Applications of Sheaves", Springer LNM **753** (1979), 479–491.

[27] P. T. Johnstone and F. E. J. Linton, Finiteness and Decidability II, Math. Proc. Camb. Philos. Soc. **84** (1978), 207–218.

[28] Powell, William C. Extending Gödel's negative interpretation to **Z**. *Journal of Symbolic Logic* **40** (1975) pp. 221–9.

[29] E.P. Specker Dualität, English translation at `http://www.dpmms.cam.ac.uk/~tf/dualityquinevolume.pdf`

23 A constructive predicative fragment of Quine's NF interprets Heyting Arithmetic

*i*NF is the constructive fragment of NF. NFP is a predicative fragment defined by blah. It is known that NFP interprets Peano arithmetic but the demonstration of this fact requires excluded middle. We show here that the constructive fragment of NFP (which we are here going to call *i*NFP) interprets Heyting Arithmetic.

The proof procedes as follows. NFP proves that V is not finite. Classically the existence of a set that is not finite is enough to furnish an interpretation of Peano arithmetic. Working in *i*NFP we reason: suppose V is finite in Kuratowski's sense (see below) then Ω (the truth-value algebra) is Kuratowski finite as well, being a quotient of a Kuratowski-finite set; but if Ω is Kuratowski-finite it has precisely two elements, so the logic is classical, so we have all the resources of NFP, and we can prove that V was not finite after all.

We can interpret Heyting arithmetic in *i*NFP if we can show that every Nfinite cardinal (see definition below) has a successor, so it will suffice to show that every Kuratowski-finite set has inhabited complement. To put it another way, it will suffice to show that there cannot be an Nfinite set whose double complement is V. It transpires that if there were such a set then the logic would be sufficiently classical to recover the proof of the axiom of infinity, which of course implies that there cannot be such an Nfinite set.

We start with definitions of Kfinite and Nfinite; kfinite $\rightarrow$ notnot Nfinite

Then we prove that if Ω is kfinite then excluded middle follows for stratified formulae. So *i*NFP proves that V is not kfinite and a fortiori not Nfinite

So no Nfinite set contains all sets.

We have to show that the complement of an Nfinite set is inhabited.

The first step is to show that the complement of any Nfinite set is nonempty; that is to say, there are no dense Nfinite sets.

For any dense Nfinite set X, the intersection of all dense Nfinite sets is the same as the intersection of all dense Nfinite subsets of X. [state this properly]

So fix such an X. We will show that the intersection of all its dense Nfinite subsets is dense.

If there are any dense subfinite sets, the intersection of all of them is also dense and subfinite as we will now show. For a start $\neg\neg$ and $\wedge$ commute so the intersection of two dense sets is dense and the intersection of an Nfinite family of dense sets is dense, by induction. The problem is that an intersection of two Nfinite sets is not obviously Nfinite.

If X and Y are two dense Nfinite sets then $x \cap Y$ is dense and subfinite is a subset of X so the intersection of all dense subfinite sets is the same as the intersection of all dense subfinite sets of blah.

Fix X a dense Nfinite set, and consider $\{X \cap Y : Y \text{ is dense Nfinite}\}$. We want (i) its intersection to be the intersection of all dense Nfinite sets and (ii) it to be dense Nfinite (or at least subfinite) as well.

(i) is easy.

(ii) $\{X \cap Y : Y$ is dense Nfinite$\}$ is a subset of the set of

Humph

The collection of kfinite subsets of X is kfinite by lemma 54, so the collection $\mathcal{X}$ of *dense* Nfinite subsets of X is subfinite. Let y be arbitrary. We know that $(\forall x \in \mathcal{X})\neg\neg(y \in x)$. Now $\mathcal{X}$ is subfinite, so we can invoke Linton-Johnstone (theorem 42) to infer $\neg\neg(\forall x \in \mathcal{X})(y \in x)$, which is to say $(\forall y)\neg\neg(y \in \bigcap \mathcal{X})$, or $\bigcap \mathcal{X}$ is dense. Let us call this minimum object '$\mathcal{V}$'. For $x \in \mathcal{V}$ and x' notnot equal to x consider $(\mathcal{V} \setminus \{x\}) \cup \{x'\}$. This, too, is dense Nfinite. Why?

LEMMA 99 *If X is Nfinite, and $x \in X$ then $X \setminus \{x\}$ is also Nfinite*

Proof: By induction on X. True when $X = \emptyset$. Now suppose true for X and suppose $y \notin X \wedge x \in (X \cup \{y\})$. We want $(X \cup \{y\}) \setminus \{x\}$ to be Nfinite. Now $y \notin X \wedge x \in (X \cup \{y\})$ implies $x \in X$. So $X \setminus \{x\}$ is Nfinite by inductive hypothesis, so $(X \setminus \{x\}) \cup \{y\}$ is Nfinite, and this is the same as $(X \cup \{y\}) \setminus \{x\}$. ■

So $(\mathcal{V} \setminus \{x\}) \cup \{x'\}$ is Nfinite. To show that it is dense it will be sufficient to show that everything in $\mathcal{V}$ is notnot in it.

$\mathcal{V} = (\mathcal{V} \setminus \{x\}) \cup \{x\}$. (This looks completely trivial but it isn't: $X = (X \setminus \{x\} \cup \{x\})$ isn't reliably true (altho' the $R \subseteq L$ inclusion is constructively correct the $L \subseteq R$ inclusion is not) but it works if X is Nfinite.) Everything in $(\mathcal{V} \setminus \{x\}) \cup \{x\}$ is either in $\mathcal{V} \setminus \{x\}$ — in which case it is certainly in $(\mathcal{V} \setminus \{x\}) \cup \{x'\}$— or it is x, and x is certainly notnot in $\mathcal{V} = (\mathcal{V} \setminus \{x\}) \cup \{x'\}$. So we must have $\mathcal{V} \subseteq (\mathcal{V} \setminus \{x\}) \cup \{x'\}$. In particular we must have $x \in (\mathcal{V} \setminus \{x\}) \cup \{x'\}$. Clearly we can't have $x \in \mathcal{V} \setminus \{x\}$ so we must have $x = x'$. So we have proved:

$$(\forall x \in \mathcal{V})(\forall y)(\neg\neg(y = x) \rightarrow y = x).$$

Now every x is notnotin $\mathcal{V}$, so every x satisfies $\neg\neg(\forall y)(\neg\neg(y = x) \rightarrow y = x)$. So we have proved

$$(\forall x)\neg\neg(\forall y)(\neg\neg(y = x) \rightarrow y = x).$$

We seek a term t such that we can prove $\neg(\forall y)(\neg\neg(y = t) \rightarrow y = t)$.

Let p be any stratified expression. We consider $\{x : p\}$ and $\{x : \neg\neg p\}$. We prove easily that they are notnotequal, but if they are actually equal we infer $p \longleftrightarrow \neg\neg p$ by extensionality as follows:

$\{x : p\} = \{x : \neg\neg p\}$ iff

$(forally)(y \in \{x : p\} \longleftrightarrow y \in \{x : \neg\neg p\})$ iff

$p \longleftrightarrow \neg\neg p$

So let t be $\{x : p\}$, and suppose $(\forall y)(\neg\neg(y = \{x : p\}) \rightarrow y = \{x : p\})$. Then, in particular $\neg\neg(\{x : \neg\neg p\} = \{x : p\}) \rightarrow \{x : \neg\neg p\} = \{x : p\})$. Now certainly $\neg\neg(\{x : \neg\neg p\} = \{x : p\})$ so we infer $\{x : \neg\neg p\} = \{x : p\})$. giving us $\neg\neg p \longleftrightarrow p$. But this is a logical principle that contradicts the existence of $\mathcal{V}$.

H I A T U S

Let Ω be the power set of the singleton of the empty set, AKA the truth-value algebra. If the logic is not classical then the following is not true:

$(\forall x)(\neg\neg(x = \{\emptyset\}) \to x = \emptyset)$

so we actually have

$\neg(\forall x)(\neg\neg(x = \{\emptyset\}) \to x = \emptyset)$,

do we not..? And this contradicts

$(\forall y)\neg\neg(\forall x)(\neg\neg(x = y) \to x = y)$ — (A)

So principle (A) is a nontrivial principle of classical logic..?

But that just says that (A) implies that the logic is notnot classical. Which is true.

H I A T U S

This means we have established that $(\forall x)(\mathrm{Nfin}(x) \to \neg\neg(\exists y)(y \notin x))$.

But what we actually wanted was $(\forall x)(\mathrm{Nfin}(x) \to (\exists y)(y \notin x))$.

If there are no dense Nfinite sets then we can establish that every (equipollence-class–style) Nfinite cardinal has a nonempty successor, as follows. If x is an Nfinite set then $\neg(\forall y)\neg\neg(y \in x)$, which is equivalent to $\neg\neg(\exists y)(y \notin x)$.

Let n be an equipollence class of Nfinite sets. Consider the set $\{X \cup \{x\} : x \notin X \in n\}$. If we can find $x \notin X \in n$ then this set is inhabited. However, by the foregoing, for any $X \in n$ we know that notnot there is such an x, so we can infer notnot of the conclusion, which is to say that $\{X \cup \{x\} : x \notin X \in n\}$ is nonempty.

Did we really need n to be inhabited? No, it was sufficient for it to be nonempty. If n is nonempty then it is notnotinhabited, so $n + 1$ is notnot nonempty, which is to say nonempty. So we seem to have proved:

If n is nonempty so is $n + 1 = \{X \cup \{x\} : x \notin X \in n\}$.

Nonempty is all very well but we want inhabited. I think we can now prove that every Nfinite cardinal is inhabited by an initial segment of the Nfinite cardinals, and that they are all distinct. Certainly $\{m : m < n\}$ is Nfinite if m is. Consider any member of m. It is Nfinite, so the set of its Nfinite subsets is Nfinite by lemma ??, and the set of cardinals thereof (which is $\{m : m < n\}$) is a quotient of an Nfinite set and is therefore Nfinite by lemma ??. (Every kfinite subset of an Nfinite set is Nfinite).

We showed above that $n + 1$ is nonempty if n is. But can we show they are distinct? Clearly if they are inhabited they are distinct. So if they are notnot inhabited they are notnotdistinct, which is to say distinct. So that's OK.

Distinctness: we prove that every equipollence number is an antichain. Then it's easy to prove they are all distinct.

H I A T U S

SO!! How does it go?

First we prove that V is not Kfinite.

Then we assume that there is a dense Nfinite set. This contradicts various weak classical principles.

If there is a dense Nfinite set then there is a least one. If there is a least one them we get certain weak classical principles. This contradicts our assumption that there is a dense Nfinite set. So there is no dense Nfinite set.

So every natural number has a successor. Done