Analogies between Jensen's theory NFU with the negation of infinity and the alternative set theory of Vopenka

Randall Holmes, Boise State University

October 20, 2025

Abstract

The subject of the talk will be formal analogies (and differences) between the alternative set theory of Vopenka on one hand and the variation of Quine's New Foundations due to Jensen augmented with the negation of the Axiom of Infinity, on the other. Both of these are theories which in some sense say that their worlds are finite, but which in fact have somewhat larger worlds. Both are theories of relatively low consistency strength with considerable expressive power for purposes of ordinary mathematics. In both cases, the applicability to normal mathematics involves some sort of thinking along the lines of nonstandard analysis, though neither theory really is exactly a system of nonstandard analysis.

Why I accepted the invitation...

I am aware of the alternative set theory of Vopenka and I have written about it, but never seriously worked in it. AST is related to nonstandard analysis, and perhaps motivated by it, but it isn't a system of NSA: it is a pure axiomatic set theory with general foundational intent, which happens to have consequences which bear on NSA. I know that AST did grow out of thoughts about nonstandard analysis, but it isn't NSA on the face of it.

I am an expert on Quine's NF and related systems. And there is a theory in this family which has some of the same features, namely, Jensen's system NFU with the additional assumption of the negation of the axiom of infinity. Like AST, this is a theory which says that all sets are finite (though some of them aren't) and which proves the existence of non-standard natural numbers, and would probably support some kind of NSA reasoning, though there is no published work of this kind (I have thought about it in the privacy of my study).

An aside

There is something I wonder about, which I cannot ask about in the most natural way, because Jensen has just died. His thesis was on subsystems of arithmetic. He published the paper on the consistency of NFU just a couple of years later. This makes me wonder whether NFU + the negation of Infinity may have been what he addressed first.

I am looking for a copy of his thesis to investigate this question.

A review of AST, possibly from an unusual standpoint

I'm going to present my own understanding of the axiomatics of AST, to facilitate communication. Odd things about what I do may be clarified by this review. I have read the axioms and thought about the theory, but I have not read much of its literature.

I'll present AST as a first-order one sorted theory. I know this isn't the usual approach but the difference is purely technical.

General objects are called *classes*. The primitive relations of the theory are equality and membership.

A *set* is defined as a class which is an element of another class:

$$\mathrm{set}(x) \leftrightarrow (\exists A : x \in A).$$

As usual, a *proper class* is a class which is not a set. A *semiset* is a proper class which is a subclass of a set. The subclass relation is defined as usual:

$$A \subseteq B \leftrightarrow (\forall x : x \in A \rightarrow x \in B).$$

The identity criterion for classes is that they have the same elements:

$$(\forall AB : A = B \leftrightarrow (\forall x : x \in A \leftrightarrow x \in B)).$$

We provide the universal class V:

$$(\forall A : x \in A \rightarrow x \in V).$$

V is the class of all sets.

The general existence principle for classes is that any property of sets determines a class. Formally, for any formula ϕ in which A does not appear free, $(\exists A : (\forall x : x \in A \leftrightarrow (x \in V \wedge \phi)))$ is a comprehension axiom, and the witness, which is unique by our identity criterion, is denoted by $\{x \in V : \phi\}$.

We have minimal set existence principles.

$\emptyset$ is defined as $\{x \in V : x \neq x\}$. We provide that $\emptyset$ is a set: $\emptyset \in V$.

If x is a class and y is a set, we define $x \cup \{y\}$ as $\{z \in V : z \in x \vee z \in y\}$. We call this the adjunction of y to x, and we provide that if x and y are sets, the adjunction of y to x is a set: $(\forall xy \in V : x \cup \{y\} \in V)$.

These are all of our set existence principles!

For any set x, we define $\{x\}$ as the adjunction of x to $\emptyset$. Of course $\{x\} = \{y \in V : y = x\}$.

We define $\{x, y\}$ (for sets x, y) as the adjunction of y to $\{x\}$. More generally, we define $\{x_1, \ldots, x_n\}$ as the adjunction of x_n to $\{x_1, \ldots, x_{n-1}\}$.

We define (x, y) as $\{\{x\}, \{x, y\}\}$ (for sets x, y) and note that this gives us an ordered pair, the proofs that it has the expected properties being familiar.

Notice that the existence of a set ordered pair lets us implement class relations quite freely.

A formula ϕ is *bounded* iff every quantifier in it is restricted to V.

The **axiom scheme of induction** asserts that for each bounded formula ϕ,
letting $I = \{x \in V : \phi\}$,

$$\emptyset \in I \wedge (\forall xy : x \in I \wedge y \in I \to x \cup \{y\} \in I) \to I = V$$

This appears to say that our universe is the collection of hereditarily finite pure sets, usually called V_ω. But it doesn't: do notice the restriction to bounded formulas. But it does say that any class definable by a bounded property which contains the empty set and is closed under adjunction is the entire universe of sets: this suggests that the universe of sets looks very much like V_ω.

An axiom of regularity is often listed, but I believe it is a consequence of induction. Nothing in my treatment hinges on this.

Nothing we have done so far is unfamiliar, though it might seem we have restricted ourselves to the hereditarily finite pure sets.

We say that a class x is *finite* iff it has no subclass which is a proper class. Note that the definition implies immediately that a finite class is a set.

The class of finite sets contains the empty set and is closed under adjunction. But notice that it is defined by a formula which is not bounded, so we cannot (or at least cannot in any obvious way) use the axiom scheme of induction to show that it is the universal class.

We define a *countable* class as a class which admits a linear order with no maximum all proper initial segments of which are finite. It is straightforward to establish by induction that no set is countable. It is easy to exhibit a countable class: the class of finite von Neumann ordinals works.

We now depart from familiar mathematics with the

axiom of prolongation: Every countable function can be extended to a set function.

This has the immediate corollary that there are semisets, since a set function extending a countable function has the countable function (a proper class) as a subclass.

The next axiom is quite technical. A pair (K, S) of a class K and a class relation S (pairing of classes can be supported as disjoint union, I note) can be taken to code a collection of classes, namely the classes $S``\{k\}$ for $k \in K$.

We say that such a code is *extensional* iff for all $k, l \in K$, $S``\{k\} = S``\{l\} \rightarrow k = l$.

The axiom of extensional coding asserts that for any code (K, S), there is an extensional code (K', S) such that $K' \subseteq K$ and (K', S) is extensional.

We note that the existence of a class well ordering of the universe (which *is* a theorem of the full theory AST) implies the axiom of extensional coding, which does seen to be a choice principle. I believe that extensional coding is needed to establish the existence of the class well-ordering: I don't think there is a redundancy here, though I wouldn't be entirely surprised if there were.

The final axiom is very simple to state.

We say as usual that two classes are the same size iff there is a class bijection between them. It is worth noting that there is a corresponding notion of two sets being the same size iff there is a *set* bijection between them: these notions do not coincide even on sets!

The final axiom asserts that

Axiom of uncountable classes: All infinite classes which are uncountable are of the same size.

I believe Vopenka thought of this as a way to simplify (and eliminate from the forefront of set theoretic thought) questions about the cardinality structure of the higher infinite.

We briefly show that the size of the universal class (the uncountable class cardinal) is ω_1.

We define a code (K, S) where S contains (x, y) if and only if the restrictions of x and y to pairs of finite von Neumann ordinals determine well-orderings of the same order type, letting K initially be the collection of all sets whose intersection with the pairs of finite von Neumann ordinals is a well-ordering. Cut this down to an extensional code (K', S) and the elements of K' code the countable ordinals. Quite standard argumentation shows that the size of K' (which we can certainly call ω_1) is uncountable, and so is the size of the universal class V and in fact also is the size of any infinite set.

This gives us AC in the form that V can be well-ordered, and it gives us the Continuum Hypothesis.

I don't want to talk extensively about mathematics in AST: first of all, I don't know as much as others here, and second, this would be an entire talk or series of talks.

I do want to remark that while this is deliberately a theory much weaker and less extensive than Cantorian set theory, it is not weak in the ways that other alternative foundations of mathematics are. It is not constructivist or predicative. Its logic is quite usual, and impredicative principles of class formation are very much in evidence.

It has the strength of third order arithmetic, which is more than enough to do all the mathematics needed for classical physics, and it can define the familiar number systems and function spaces needed for classical physics in efficient if not entirely familiar ways.

A model construction for AST

I'm going to discuss the construction of a model of AST along lines which are quite familiar to my audience, no doubt, and no doubt are known. The final move of this section is a bit surprising, and I do wonder if it is familiar to the AST community.

We work in the usual set theory ZFC (and not in very much of this theory, which will be important for our final move). We assume that the continuum hypothesis holds (it is well known that this can be done, and I am not going to talk about that at all!)

The sets of our model are implemented using the elements of an ultrapower on V_ω. Let U be a nonprincipal ultrafilter on $\mathbb{N}$ (considered as the set of von Neumann ordinals, so as a subset of V_ω).

Let M_0 be the collection of functions $f : \mathbb{N} \to V_\omega$. For $x, y \in M_0$, we define $x =_{M_0} y$ as $\{n \in \mathbb{N} : x(n) = y(n)\} \in U$ and $x \in_{M_0} y$ as $\{n \in \mathbb{N} : x(n) \in y(n)\} \in U$. Familiar math shows that M_0 with these equality and membership relations gives a model of the full theory of V_ω.

Now the elements of our model M are all of the subclasses of M_0 which are unions of equivalence classes under $=_{M_0}$. With each element x of M_0 we naturally associate an element of the model M, called $\texttt{ext}(x)$, defined as $\{y \in M_0 : y \in_{M_0} x\}$. We define $x \in_M y$ as holding iff for some $x' \in M_0$, $x = \texttt{ext}(x')$ and $x' \in y$.

Review of the axioms

Extensionality obviously holds: if two items in M contain the same elements in the sense defined above, then they include the same equivalence classes under $=_{M_0}$ and so are equal in the metatheory.

Each equivalence class under $=_{M_0}$ is associated with the common preimage of each of its elements under $\in_{M_0}$: the latter objects are the sets in M. Clearly any collection of the sets of M in the sense of the metatheory yields a class of M (the union of the corresponding equivalence classes) so class comprehension holds.

The axioms of empty set, adjunction, and induction hold because M_0, and so the corresponding system of sets of M, is a model of the full first order theory of V_ω.

The trickiest bit is the axiom of prolongation. Suppose C is a countable function in M, with a given linear order. For each $n \in \mathbb{N}$, let c_n be a function belonging to the equivalence class in M_0 associated with the nth element of C in the linear order. Now consider the map F from $\mathbb{N}$ to V_ω which sends each natural number n to $\{c_m(n) : m < n\}$. The equivalence class of F in M_0 is a function in the sense of M_0 interpreted as a nonstandard model of V_ω, and $c_n \in_{M_0} F$ for every n: the element of M corresponding to the equivalence class of F is a set of M, is a function in M's internal sense, and extends the class function C.

Clearly the cardinality of M_0 is $c = \omega_1$, and every element of M is the union of either finitely many equivalence classes, ω equivalence classes, or ω_1 equivalence classes, simply because CH holds in the metatheory.

The axiom of extensional coding holds because there is a class well-ordering of the classes of the model.

The surprise

The surprise (to me, perhaps it is generally known) is that this construction can be carried out entirely in AST, giving a class model of AST which is in effect an ultrapower.

Of course, this requires attention to coding of classes by sets. But it is straightforward.

It is straightforward to construct a representation of a nonprincipal ultrafilter U on the natural numbers (considered as finite von Neumann ordinals) by induction along the sets (with a general set taken as coding its intersection with the class of finite von Neumann naturals) in an essentially standard way.

Then general functions from $\mathbb{N}$ to V_ω are again codable by sets (a general set coding its intersection with $\mathbb{N} \times V_\omega$ when this happens to be a function) and the relations $=_{M_0}$ and $\in_{M_0}$ are definable as class relations.

The domain of classes which are unions of equivalence classes under $=_{M_0}$ is then something we can talk about, define the interpreted membership relation on appropriately, and carry out the verification of the axioms much as above.

There is no Gödel issue here, any more than with the construction of a class model of $ZF+V=L$ in a generic model of ZF.

What this says is that we can interpret the theory of an ultrapower model of nonstandard analysis in AST and see that AST actually holds in this interpreted model.

But other things hold in this model which do not seem to hold in AST in general. I haven't made my thoughts about this precise enough, but certainly this is true. In the model of AST interpreted inside AST as above, we can ask, given a countable class C, for "elements" of C with nonstandard natural number indices in the order on C in a quite uniform manner. In this interpretation of AST, we have something like the principle of standardization in IST. This does not seem to be workable in general models of AST: does anyone here know if there are actual results about this?

Working from the other end...an account of NFU with the negation of Infinity

NFU is a first order theory with equality and membership as primitive relations. It is also usual to have a primitive constant $\emptyset$.

The axiom for $\emptyset$ asserts that $(\forall x : x \notin \emptyset)$. A *set* is an object which either has an element or is equal to $\emptyset$.

The identity criterion for sets is that they have the same elements. Note that the axioms and the definition of sethood tell us that a non-set, if there is one, has no elements.

$$(\forall xy : \mathtt{set}(x) \wedge \mathtt{set}(y) \rightarrow (\forall z : z \in x \leftrightarrow z \in y))$$

We say that a formula ϕ is stratified iff there is a function σ from variables to natural numbers such that for each atomic subformula $x = y$ of ϕ we have $\sigma(x) = \sigma(y)$ and for each atomic subformula $x \in y$ we have $\sigma(x) + 1 = \sigma(y)$.

More heat and light has been made about use of defined notions in NFU (or in NF) than is really required. For example, all definable term operations can be introduced at once by introducing a definite description operator $(\theta x : \phi)$ intended to represent the unique x such that ϕ if there is one, and otherwise $\emptyset$. The stratification condition for an atomic subformula involving a definite description is that a term $(\theta x : \phi)$ has the same image under σ as x (determining the rules for its participation in atomic subformulas). One can then observe that the natural method of eliminating $(\theta x : \phi)$ from formulas preserves stratification: nothing is really added. Relations with stratified definitions are easily handled.

Rather than try to do mathematics in the unfamiliar stratified context, we will talk about a class of models of NFU.

Work in a nonstandard model of the usual set theory with an automorphism which moves a rank of the cumulative hierarchy.

Let α be a (necessarily nonstandard) ordinal of the model such that $j(\alpha) \neq \alpha$. We can without loss of generality assume $j(\alpha) < \alpha$.

The domain of our model of NFU will be the nonstandard rank V_α. The equality relation of the model will be the restriction of the equality of the nonstandard model of set theory to V_α^2. The membership relation of the model of NFU is defined as $x \in_{\text{NFU}} y \leftrightarrow j(x) \in y \wedge y \in V_{j(\alpha)+1}$.

The idea here is that the automorphism lets us treat $V_{j(\alpha)+1}$, a rank lower than and included in V_α as the power set of V_α; objects not in $V_{j(\alpha)+1}$ (most objects!) become non-sets, with no elements.

We will not attempt to present the proof that this works in the confines of this talk. This is a standard argument (a variant of Jensen's original approach to proving the consistency of NFU). What it does for us is give us a sensible picture of what the world of NFU is like. These models have special features which do not hold in NFU in general, but general NFU supports an interpretation of NFU extended so that all these special features are present.

Notice that nothing here prevents α from being finite, or indeed prevents us from contemplating a nonstandard model of V_ω with an automorphism moving a natural number, which is in fact the approach we will take. A model built in this way will satisfy the negation of the Axiom of Infinity, though it will be covertly infinite and contain "infinite" nonstandard natural numbers, a phenomenon familiar from AST.

We proceed to construct a model of NFU using ultrapowers. What we actually do is build a nonstandard model of V_ω using ultrapowers which evidently has an automorphism moving a natural number, then appeal to our model construction above.

We will use a system of ultrapowers rather than a single one. For each $n \geq 1$, we need a nonprincipal ultrafilter on $\mathbb{N}^n$, the set of n-tuples taken from V_ω.

U_1 is a nonprincipal ultrafilter on $\mathbb{N}^1$.

U_2 is an ultrafilter extending

$\{A \subseteq \mathbb{N}^2 : \mathrm{dom}(A) \in U \wedge \mathrm{rng}(A) \in U \wedge A \subseteq [<_{\mathbb{N}}]\}.$

U_{n+1} is an ultrafilter extending

$$\{A \subseteq \mathbb{N}^{n+1} : \{a \lceil [0, n-1] : a \in A\} \in U_n \wedge$$

$$\{(a \lceil [1, n]) \circ \sigma) : a \in A\} \in U_n \wedge \mathtt{increasing}(A)\}$$

[note that we are treating an element of $\mathbb{N}^n$ as a function from $\{0, n-1\}$ to $\mathbb{N}$ for purposes of expressing this definition; σ is the successor operation, used here for reindexing].

The idea here is that the ultrafilters describe indiscernibles.

Showing that the sets in the last two clauses of the definition are filters is left as an exercise.

The idea is the the ultrafilters describe increasing sequences of indiscernible nonstandard natural numbers.

A function F from $\mathbb{Z}^\omega$ to V_ω has finite interval support iff there is an interval $[a, b]$ such that $F(v)$ is exactly determined by $v{\restriction}[a, b]$ for each $v \in \mathbb{Z}^\omega$.

The elements of our ultrapower model M_0 of V_ω are the functions F from $\mathbb{Z}^\omega$ to V_ω with finite interval support.

Define $j(F(n))$ as $F(n-1)$.

For R either of the relations $=$ or $\in$, $FR_{M_0}G$ is defined as holding iff

$$\{v \in \mathbb{N}^{b-a+1} :$$

$$(\forall v' \in \mathbb{N}^\omega :$$

$$v' \supseteq v \to j^a(F)(v')Rj^a(G)(v'))\} \in U_{b-a+1}$$

for each (or for some) common finite interval support $[a, b]$ of F and G. Note that the way the ultrafilters are defined ensures that it does not matter which finite interval support is chosen. The function j (iterated a times) is used to translate the finite interval support $[a, b]$ to the domain $[0, b-a+1]$ relevant to U_{b-a+1}.

This structure is a nonstandard model of the full theory of V_ω for standard reasons. The idea is that we have an integer-indexed sequence of indiscernibles described by the ultrafilters (U_n tells us which standard sets any contiguous sequence of length n of the indiscernibles lives in) and elements of our model are determined by standard functions acting on a contiguous finite sequence of the indiscernibles.

We note with interest particular elements of the model: $c_n(v) = v(n)$. These are natural numbers of the model, and $c_m < c_n$ holds iff $m < n$. These are our indiscernibles.

The function defined by $j(F)(n) = F(n-1)$ is an automorphism of this structure, moving c_n to c_{n-1} for each $n \in \mathbb{Z}$.

So the nonstandard rank indexed by $c(0)$ in this model can be used to define a model of NFU as above, since the automorphism j moves $c(0)$ to $c(-1)$, which is less than $c(0)$ in the sense of the model.

It can be noted that this entire procedure can be carried out in AST, though tracking down the possibility of doing all the coding involved is tricky.

Another surprise

The question which remained in my mind after doing the previous was the status of the Prolongation Axiom in NFU.

It is fairly easy to establish that in the kind of model we have just described, any countable function with range in the finite natural numbers can be extended to a set (glossing over exactly how the finite numbers are represented inside NFU; this is not difficult but would take valuable time). The reason is that there is a standard function f from $\mathbb{N} \to \mathbb{N}$ in play, and we can take any c_m (considered as a nonstandard finite number of the model) and define the restriction of f to the natural numbers less than c_m as a set of the model of NFU.

This is a kind of standardization operation. But it isn't Prolongation, because we do not allow arbitrary model elements in the range of the countable class function.

The reason that we cannot do this is that we cannot be sure that all elements of the countable class have a common interval support, and without a common interval support, we have no obvious way to construct a nonstandard finite initial segment of the function.

This is related to the known fact that in any model of NFU, there are countable proper classes. We give an example without supporting proofs. For any natural number n define Tn as the cardinal of the elementwise image of any set of cardinality n under the singleton operation. T is an operator definable in NFU; in the class of models we are considering, it happens to be the restriction of j to the natural numbers. It is provable that if n is the finite size of the universe, $Tn < n$ and in general the external sequence $T^i(n)$ is strictly decreasing. It is evident that this holds in the models we are discussing. But then the collection of $T^i(n)$'s cannot be a set, because it is a collection of natural numbers of the model with no smallest element.

But it turns out that we **can** have Prolongation in a theory extending NFU with the negation of infinity.

We extend NFU to allow proper classes. The stratified comprehension axiom for sets acquires the additional restriction that all quantifiers must be bounded in the collection V of all sets, and all parameters must be sets. This theory is called MLU for historical reasons.

Any model of NFU can be augmented to a model of this theory by implementing proper classes as arbitrary subcollections of the model which are not extensions of elements of the model (the model elements represent their own extensions, of course).

In MLU, we can state the Prolongation Axiom and the assertion that all classes have cardinality either finite, ω, or ω_1.

We obtain our model in a way which is surprising and bizarre, but which is hinted at by trying to reason about the Prolongation Axiom in MLU from first principles (we won't attempt this here).

In a metatheory in which CH holds, build a model of NFU as above. Then expand the model of NFU to an ultrapower, using a further nonprincipal ultrafilter on the finite natural numbers of the model. Over this expanded model, there are ultrafilters U_n indexed by nonstandard natural numbers. Pass to the model of MLU in which classes are arbitrary subcollections of this model.

Consider any countable class of this model. All elements of the range of the countable class will actually have a common (nonstandard) finite interval support: consider the finite interval supports $[a_i, b_i]$ of the elements C_i of the class. There will be a nonstandard a less than all the a_i's, and a nonstandard b greater than all the b_i's, because the nonstandard extension of the finite natural numbers of the initial model of NFU has cofinality ω_1. So we get a common finite interval support $[a, b]$ for all the C_i's and we can describe a nonstandard finite initial segment of the function $(i \mapsto C_i)$ in a way analogous to our demonstration of the Prolongation Axiom in our model of AST above.

This is admittedly a sketch: I discovered this while preparing the talk! It does *appear* that this construction can be carried out in AST, so the two theories are mutually interpretable.

The theory MLU + NegInf + Prolongation + the infinite cardinals of classes are ω and ω_1 realizes the concept I started with that NFU with the negation of infinity can be construed as having a world rather like that of AST. It is worth noting that NFU + NegInf itself is a much weaker theory. I am interested (was already interested) in seeing how much non-standard analysis one can do in this weak theory. I am also interested in investigating the stronger theory, which I would not initially even have thought consistent.

A couple of observations which indicate ways in which these theories are different, or seem different.

Of course NFU and MLU as extended differ from AST in having lots of atoms, that is, non-sets. This difference is superficial, but it takes some work to discover this.

MLU as extended deviates from expectations which students of NF-like theories have about unstratified extensions of the theory: the usual expectation is that we want the natural numbers of sets which are the same size as their images under the singleton operation (natural numbers which are fixed by j, in terms of our models) to be standard, and here our device for getting the Prolongation Axiom requires us to have nonstandard natural numbers fixed by j. The extension of MLU actually proves that there are natural numbers of the kind indicated which have to be nonstandard, entirely internally.