THE REVIEW OF SYMBOLIC LOGIC
Volume 8, Number 3, September 2015

NAIVE SET THEORY AND NONTRANSITIVE LOGIC

DAVID RIPLEY
Department of Philosophy, University of Connecticut

Abstract. In a recent series of papers, I and others have advanced new logical approaches to familiar paradoxes. The key to these approaches is to accept full classical logic, and to accept the principles that cause paradox, while preventing trouble by allowing a certain sort of *nontransitivity*. Earlier papers have treated paradoxes of truth and vagueness. The present paper will begin to extend the approach to deal with the familiar paradoxes arising in naive set theory, pointing out some of the promises and pitfalls of such an approach.

§1. Introduction. In a recent series of papers Cobreros *et al.* (2012), Cobreros *et al.* (2013), Ripley (2012), Ripley (2013a), and Ripley (2013b), I and others have advanced new logical approaches to familiar paradoxes. The key to these approaches is to accept full classical logic (in a certain sense), and to accept the principles that cause paradox, while preventing trouble by allowing a certain sort of *nontransitivity*. Earlier papers have treated paradoxes of truth and vagueness. The present paper will begin to extend the approach to deal with the familiar paradoxes arising in naive set theory, pointing out some of the promises and pitfalls of such an approach.

Naive set theory, for my purposes here, is a theory of sets with two ingredients: comprehension and extensionality. Comprehension guarantees that, for any formula with one variable free, there is a set that contains all and only the things that satisfy the formula. This guarantees the existence of perfectly ordinary sets, like the set of all even numbers, as well as the existence of paradoxical sets, like the set of all nonselfmembered sets (the russell set, which will play a prominent role in what follows). With comprehension in place, there is no need for many of the separate existence axioms used in ZF set theory: we automatically have the existence of pair sets, union sets, images under functions, and so on, directly from comprehension. If the language can state a condition, there is a set answering to it.

Extensionality is the principle that set identity depends only on membership: if two sets have the same members, they are in fact the same set. Without extensionality, as I see it, we do not have a *set* theory at all, but only something like a property theory. Distinct properties can be instantiated by the very same things, but distinct sets must have some difference in membership. Comprehension gets most of the press in many treatments of naive set theory, but extensionality is no less challenging. In a variety of different logical settings, extensionality can cause trouble. But here, extensionality will not be so difficult.

The set theory to be described in this paper will start from a formulation of first-order classical logic, and will add comprehension and extensionality. The resulting system will allow for proofs of the existence of sets of all sorts, including paradoxical sets. Nonetheless, as I will show, the theory is nontrivial: many things cannot be proven in it. (In fact, I will show something stronger—quantifier-free model-theoretic conservativity—from which nontriviality is a quick corollary.)

Received: August 28, 2014.

doi:10.1017/S1755020314000501

The remainder of the introduction will present a sequent system for classical logic; it is this sequent system that will serve as the base logic for our set theory. The system is to some extent arbitrary; many different presentations could work for the same purpose. I'll call attention to those features that matter. §§2 considers three ways of adding comprehension to our sequent system. I will take a goldilocks approach, and argue that two of them miss the mark: one is too weak and one is too strong. The third is just right. Similarly, §§3 considers three ways of adding extensionality. Again, one is too weak and one is too strong, while the third is just right. Once I have presented the final formulations of both comprehension and extensionality, we will be ready to consider models for the system; §§4 does just this. As I will show there, this set theory has much in common with naive set theories built on the paraconsistent logic LP, despite the very different way I'll arrive at it. I will use the models presented there to demonstrate the nontriviality of the system.

1.1. The base logic. I'll work with a standard first-order language with equality, including set-abstract terms $\{x : A\}$ for every formula A and variable x, as well as the more usual constants and variables, and a distinguished binary predicate $\in$ for membership.

One goal of the nontransitive approach I'll be pursuing here is to validate all arguments that are valid in classical logic. This will be achieved by working within a sequent formulation of classical logic with identity; the additional rules for handling naive sets will only add to this system. The system I'll use is given in Figure 1; I'll call it CL=, to have a name for it.[1] Most of the precise details of CL= are not very important; there are many other possible formulations of classical logic that would work just as well.

Some of its details are very important, though. Cut is the following rule, which is quite crucially *not* a rule of CL=:

$$\frac{[\Gamma : A, \Delta] \qquad [\Gamma', A : \Delta']}{[\Gamma, \Gamma' : \Delta, \Delta']}$$

Cut encodes a certain form of *transitivity* of the consequence relation in question. Although cut is not a rule of CL=, it is admissible. That is, where CL=$^+$ is the system that results from adding the rule of cut to CL=, any sequent derivable in CL=$^+$ is already derivable in CL=. So I'm not depriving us of any validities by avoiding cut. But I am leaving myself room I would not otherwise have—room to extend the system in novel directions. More later.

The other feature of note in this sequent calculus is the two drop rules: $\bot$-drop and =-drop. These are slightly different from more usual formulations, which use axioms $[\bot : \ ]$ and $[\ : t = t]$. In the presence of the axiom $[A : A]$, these usual axioms can be quickly derived using the drop rules. On the other hand, deriving the drop rules from the usual axioms requires use of a cut rule. In the absence of cut, then, the drop rules are stronger than the usual axioms.[2]

---

[1] In a sequent $[\Gamma : \Delta]$, Γ and Δ are finite sets of wffs; by using sets we avoid any need for structural rules of contraction or exchange. As usual, I abbreviate things like $[\Gamma \cup \{A\} : \emptyset]$ as $[\Gamma, A : \ ]$, and so on. Further, in this figure, A/B can be either A or B, t can be any term whatever, and a is an *eigenvariable*: a variable that does not occur free in the conclusion-sequent of its rule.

[2] Since cut is admissible (for now), there is no immediate upshot to this difference. The difference will reveal itself in a moment, when the system is extended to one that does not admit cut. This drop-rule strategy for handling axioms without cuts is taken from Negri & von Plato (1998). These drop-rule formulations are equivalent to the usual axiom formulations together with a highly restricted rule of cut, which allows cuts only on $\bot$ or $t = t$.

Axioms:

$$\text{Id:}\quad \frac{}{[A : A]}$$

Structural rules:

$$\text{KL:}\quad \frac{[\Gamma : \Delta]}{[\Gamma, A : \Delta]} \qquad \text{KR:}\quad \frac{[\Gamma : \Delta]}{[\Gamma : A, \Delta]}$$

Operational rules:

$$\bot\text{-drop:}\quad \frac{[\Gamma : \Delta, \bot]}{[\Gamma : \Delta]}$$

$$\neg\text{L:}\quad \frac{[\Gamma : A, \Delta]}{[\Gamma, \neg A : \Delta]} \qquad \neg\text{R:}\quad \frac{[\Gamma, A : \Delta]}{[\Gamma : \neg A, \Delta]}$$

$$\wedge\text{L:}\quad \frac{[\Gamma, A/B : \Delta]}{[\Gamma, A \wedge B : \Delta]} \qquad \wedge\text{R:}\quad \frac{[\Gamma : A, \Delta] \qquad [\Gamma : B, \Delta]}{[\Gamma : A \wedge B, \Delta]}$$

$$\vee\text{L:}\quad \frac{[\Gamma, A : \Delta] \qquad [\Gamma, B : \Delta]}{[\Gamma, A \vee B : \Delta]} \qquad \vee\text{R:}\quad \frac{[\Gamma : A/B, \Delta]}{[\Gamma : A \vee B, \Delta]}$$

$$\supset\text{L:}\quad \frac{[\Gamma : A, \Delta] \qquad [\Gamma', B : \Delta']}{[\Gamma, \Gamma', A \supset B : \Delta, \Delta']} \qquad \supset\text{R:}\quad \frac{[\Gamma, A : B, \Delta]}{[\Gamma : A \supset B, \Delta]}$$

$$\forall\text{L:}\quad \frac{[\Gamma, A : \Delta]}{[\Gamma, \forall x A[x/t] : \Delta]} \qquad \forall\text{R:}\quad \frac{[\Gamma : A, \Delta]}{[\Gamma : \forall x A[x/a], \Delta]}$$

$$\exists\text{L:}\quad \frac{[\Gamma, A : \Delta]}{[\Gamma, \exists x A[x/a] : \Delta]} \qquad \exists\text{R:}\quad \frac{[\Gamma : A, \Delta]}{[\Gamma : \exists x A[x/t], \Delta]}$$

= rules:

$$\text{=L1:}\quad \frac{[\Gamma, A : \Delta]}{[\Gamma, t = u/u = t, A[u/t] : \Delta]} \qquad \text{=L2:}\quad \frac{[\Gamma : A, \Delta]}{[\Gamma, t = u/u = t : A[u/t], \Delta]}$$

$$\text{=-drop:}\quad \frac{[\Gamma, t = t : \Delta]}{[\Gamma : \Delta]}$$

Fig. 1. The calculus CL=.

It can be shown by the usual techniques that this calculus is sound and complete for familiar classical models; this is full classical logic. Moreover, it will remain fully in force throughout the paper; no axiom or rule will be weakened or given up. The goal is to treat naive set theory by *adding to* classical logic, rather than by taking anything away.

But—and this is the fun part—we can add selectively, if we choose. Although the given system admits cut, the same is not true of all its extensions. It's possible to add validities *without* adding everything that would follow from those validities via cut. By doing this, it's possible to arrive at various nontransitive systems stronger than classical logic.

Moreover, these nontransitive systems can be quite interesting for the treatment of various befuddling phenomena. This is just the tactic I and others have used elsewhere (eg Ripley, 2013a, 2013b; Cobreros *et al.*, 2012) to provide treatments of familiar paradoxes of truth and vagueness that manage both to accept strong intuitive principles (transparent truth, tolerant vagueness) and to preserve the validity of every classically-valid argument. It is the tactic I'll use in this paper as well, to explore the possibilities for adding naive comprehension and extensionality to CL=. The resulting system will show that the approach has promise for the paradoxes of naive set theory as well.

§2. Comprehension. This section will consider three different ways to add naive comprehension to CL=. The first two are more traditional, and I will show that neither of them will work. One is far too weak, and the other far too strong. The third approach may be a bit less familiar, but it is just right.

2.1. As an axiom. I'll call $\exists y \forall x(x \in y \equiv A)$, with y not free in A, the *naive comprehension schema*, or NC. A number of other approaches to naive set theory (Brady, 1989; Priest, 2006; Restall, 1992; Routley, 1977; Weber, 2012) fix on NC to express their naivete.[3] And indeed, we should expect all instances of NC to be theorems of any naive set theory: a failure of NC would indicate a failure of full comprehension.

We might try to be as straightforward as possible about this, and simply add to CL= the axioms $[\;:A]$, for any instance A of NC. Call the resulting system NC1. This has one pleasant (and immediate!) effect: all instances of NC are theorems of NC1.

Moreover, the paradoxes are handled in just the way I'm going for: cut is no longer admissible in NC1. To see this, consider the russell set. Since the base system CL= is classical, there is a cutfree derivation in NC1 of $[a \in a \equiv a \notin a :\;]$. Starting from there:

$$\forall\text{L:}\;\frac{[a \in a \equiv a \notin a :\;]}{[\forall x(x \in a \equiv x \notin x) :\;]} \qquad \exists\text{L:}\;\frac{[\forall x(x \in a \equiv x \notin x) :\;]}{[\exists y \forall x(x \in y \equiv x \notin x) :\;]}$$

In NC1, we also have as an axiom $[\;: \exists y \forall x(x \in y \equiv x \notin x)]$. Thus, if cut were admissible in NC1, the empty sequent would be derivable. But the empty sequent is not derivable in NC1.[4] So NC1 includes the full naive comprehension schema, and succeeds in handling the resulting paradoxes by failing to admit cut. These are just the sort of features I'm looking for here. But still, NC1 will not do; it is far too weak.

For example, there is no derivation in NC1 of the sequent $[\;: \exists y(t \in y \supset Pt)]$.[5] But this ought to be a very direct consequence of naivete; just let y be the set of Ps. Then if t is in y, it must indeed be P. If something this simple is not a theorem of NC1, there is little hope of being able to use NC1 to work with sets in any useful way. We need something stronger.

2.2. As a drop rule. A natural second thought is to stick with the NC schema, but to impose it via a drop rule of the sort used in CL= for $\bot$ and $t = t$. That is, we might add the following rule, for every instance A of NC:

$$\text{NC-drop:}\quad \frac{[\Gamma, A : \Delta]}{[\Gamma : \Delta]}$$

[3] Of course, different logics have different conditionals, and most of these authors have used distinctive conditionals in their formulations of NC. Here, I stick to the material conditional without considering other options, which would take me too far afield.

[4] NC1 is far weaker than my eventual target system; it will be proven in §4.2 that this target system includes no derivation of the empty sequent. Alternately, see footnote 5 for a more direct strategy.

[5] Due to the presence of the drop-rules, NC1 doesn't have the subformula property. But it has the almost-subformula property: in any derivation, all formulas that appear are either 1) subformulas of formulas appearing in the conclusion sequent, or 2) $\bot$, or 3) $t = t$, for some term t. Now it's quick to show there's no derivation of $[\;: \exists y(t \in y \supset Pt)]$. NC itself doesn't meet any of the three conditions in this case, so any derivation can't have involved the NC axiom; thus, the sequent is derivable in NC1 iff it's derivable in CL=. But the consequence relation determined by CL= is just the familiar classical consequence relation, and this sequent clearly isn't classically valid.

In words, if a sequent is derivable with an instance of NC as a premise, then the instance can be dropped; the sequent is derivable without the instance as well. Let NC2 be CL= plus NC-drop. As I mentioned above, in the absence of cut a drop rule like this is strictly stronger than its corresponding axiom. As such, one might hope that NC2 gets more than NC1 did.

This is, alas, too right. NC2 is trivial. There is a derivation of the empty sequent, and via the K rules this allows for any sequent at all to be derived. One way of doing this goes via the familiar russell set. Above, we saw that $[\exists y \forall x(x \in y \equiv x \notin x) : \ ]$ is derivable in in CL=. But then it is only one step to disaster:

$$\text{NC-drop:} \quad \frac{[\exists y \forall x(x \in y \equiv x \notin x) : \ ]}{[\ : \ ]}$$

So NC2 is hopeless.

NC, then, is not the right approach to take to comprehension in this setting. One can either have NC imposed via an axiom (as in NC1), which is all well and good but not enough, or have it imposed via a drop rule (as in NC2), which is overstrong and trivializes the resulting system. These failures lead me to turn away from NC-based approaches to naive comprehension altogether.

2.3. The comprehension rule. There is another option, however. Rather than trusting in NC to do the work, there's a way to establish the connection between 't is P' and 't is a member of the set of Ps' more directly. The goal is to ensure that $A(t)$ and $t \in \{x : A(x)\}$ should be interchangeable; this can be established with a pair of (two-way!) rules:

$$\text{CompL:} \quad \frac{[\Gamma, A(t) : \Delta]}{\overline{[\Gamma, t \in \{x : A(x)\} : \Delta]}} \qquad \text{CompR:} \quad \frac{[\Gamma : A(t), \Delta]}{\overline{[\Gamma : t \in \{x : A(x)\}, \Delta]}}$$

Let CL=Comp be CL= plus these rules.[6] So $A(t)$ and $t \in \{x : A(x)\}$ are interchangeable in CL=Comp, as premises or as conclusions. It follows from this, by induction on formula construction, that they are interchangeable as subformulas of premises and conclusions as well.[7]

CL=Comp, then, has one of the key features that motivated naive set theory in the first place: predications of all sorts are fully interchangeable with the corresponding membership statements. There is no difference, in this system, between the claim that t is P and the claim that t is a member of the set of Ps—except that the latter claim includes a term 'the set of Ps' to be quantified over, substituted for, etc.[8]

---

[6] In the use of these rules from top to bottom, there is no assumption that t is free in $A(t)$, nor is there any assumption that every free occurrence of t is replaced with x. For example, if $A(t)$ is $t = t$, then $A(x)$ can be $t = t$, $t = x$, $x = t$, or $x = x$. That is, t's self-identity is interchangeable with its membership in each of the following: 1) the set of things such that t is self-identical, 2) the set of things that t is identical to, 3) the set of things that are identical to t, and 4) the set of things that are self-identical. On the other hand, in the use from bottom to top, we must require that every free occurrence of x in $A(x)$ is replaced with t to yield $A(t)$; otherwise we might end up with newly-free xs floating about.

[7] There is another, slightly weaker, natural option here; I will discuss it briefly in §4.1. But the Comp rules above will be my main focus.

[8] An anonymous referee worries that this goes too far, that 'the informal idea behind naive comprehension' only requires NC or some relative, rather than this full intersubstitutability. While I cannot give the issue the full discussion it deserves here, the situation seems to me strikingly analogous to one in the theory of naive truth; what I'm offering here can be seen as something

This is enough to get all the benefits of NC1, limited as they were. Recall that $[A : A]$ is derivable, for any formula A. Then we have:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[A(x) : A(x)]}{[x \in \{y : A(y)\} : A(x)]}\,\text{CompL}}{[\ : x \in \{y : A(y)\} \supset A(x)]}\,{\supset}\text{R} \qquad \dfrac{\dfrac{[A(x) : A(x)]}{[A(x) : x \in \{y : A(y)\}]}\,\text{CompR}}{[\ : A(x) \supset x \in \{y : A(y)\}]}\,{\supset}\text{R}}{[\ : A(x) \equiv x \in \{y : A(y)\}]}\,\wedge\text{R}}{[\ : \forall x(A(x) \equiv x \in \{y : A(y)\}]}\,\forall\text{R}}{[\ : \exists y \forall x(A(x) \equiv x \in \{y : A(y)\}]}\,\exists\text{R}$$

This shows that CompL and CompR suffice to render every instance of NC a theorem of CL=Comp. So CL=Comp is at least as strong as NC1; there is no need to take the instances of NC as axioms.

But CL=Comp in fact goes well beyond NC1. Recall the sequent $[\ : \exists y(t \in y \supset Pt)]$. Where NC1 could not derive this sequent, CL=Comp derives it in just the way you'd expect:

$$\dfrac{\dfrac{\dfrac{[Pt : Pt]}{[t \in \{z : Pz\} : Pt]}\,\text{CompR}}{[\ : t \in \{z : Pz\} \supset Pt]}\,{\supset}\text{R}}{[\ : \exists y(t \in y \supset Pt)]}\,\exists\text{R}$$

CL=Comp also avoids the triviality that infects NC2, as will be shown in §4.2. So CL=Comp sails between the Scylla of NC1 and the Charybdis of NC2; its naivete is strong enough to demonstrate the sort of results that a naive set theory ought to include, but not so strong as to trivialize the system. Since CL=Comp includes NC1, it must be nontransitive to be nontrivial, just as NC1 was. So it embodies the strategy I'm aiming for.

§3. Extensionality. CL=Comp, I think, is a fine theory of naive properties. It allows for a fully naive treatment of all sorts of paradoxical properties by adding naive comprehension to a cutfree presentation of classical logic and allowing transitivity to fail.

But there is no need to stop there. As I mentioned in the introduction, the goal here is to give a full naive *set* theory. Sets, though, are unlike properties, in that their identity is determined by their extensions: sets with the same members must be the same set. Nothing in comprehension guarantees that. It is a separate demand.

This is where a variety of theories, naive and otherwise, run into trouble. For example, in Łukasiewicz's infinite-valued logic, naive comprehension on its own causes no trouble (White, 1979), but the addition of extensionality trivializes the resulting system (allows any sentence at all to be proved) (Restall, 1994, p. 225). In a quite different setting, Hinnion & Libert (2003) shows that extensionality is sufficient to trivialize in a classical setting together with only a very limited form of comprehension. Restall (2013) develops this into a quite general challenge to naive set theories of various stripes.[9] Another example: Grišin

like a deflationist approach to set membership, although one quite unlike that of Incurvati (2012). Just as in the case of truth, deflationism can motivate naivete, but naivete itself does not seem to require deflationism. See also Beall (2009), Glanzberg (2005), and Shapiro (2011).

9 The set theory of this paper provides an answer to Restall's challenge, which I take to be deep and serious. In brief: Restall shows that any nontrivial set theory must reject at least one of a very small list of very plausible-looking principles. Standard sophisticated set theories, for example,

(1982) develops a contraction-free theory for naive comprehension that trivializes on the addition of extensionality. In yet another setting, Field (2008) offers a consistent naive property theory, but as is shown in Field *et al.* (2014), this logic (and its successor in Field, 2014) cannot accommodate extensionality.

Finally, even where extensionality can safely be added, it is often a lot of work. For example, Brady (2006) lays out a naive set theory, but must devote an entire chapter—and the hardest part of the proof—to showing that extensionality causes no trouble.

So extensionality is very much not just to be had for free. But it can be had. Here, I will consider three ways of adding extensionality to CL=Comp. Again, one will turn out to be too weak, a second too strong, and the third just right. Moreover, the eventual target formulation of extensionality fits into the nontriviality proof in a particularly elegant and simple way; because of this, it may have applications in nonstandard set theories more generally.

3.1. As an axiom. Consider $\forall x \forall y(\forall z(z \in x \equiv z \in y) \supset x = y)$, which I'll call E. This sentence is a natural statement of extensionality. At the very least, it ought to be a theorem of a naive set theory.

We could opt for the straightforward approach, and simply take $[\ : \mathrm{E}]$ as an axiom. Let E1 be CL=Comp plus this axiom. E1, it turns out, is again far too weak to be helpful. It does not allow even simple sequents like $[\ : (\forall x(Px \equiv Qx)) \supset \{x : Px\} = \{x : Qx\}]$ to be derived.[10] But this should surely be a theorem of a properly extensional set theory.

3.2. As a drop rule. This might lead us to reach for a drop-rule formulation.

$$\text{E-drop:}\quad \frac{[\Gamma, E : \Delta]}{[\Gamma : \Delta]}$$

Let E2 be CL=Comp plus E-drop. Unfortunately, E2 is trivial; there is a derivation of the empty sequent. In what follows, let h be the empty set $\{x : \bot\}$, let r be the russell set $\{x : x \notin x\}$, and let w be the *weber set* $\{x : r \in r\}$.[11]

I'll proceed in stages, for readability. First, derive $[r \in r :\ ]$:

$$\begin{array}{r c}
\neg\text{L:} & \dfrac{[r \in r : r \in r]}{[r \in r, r \notin r :\ ]} \\
\text{CompL:} & \dfrac{}{[r \in r :\ ]}
\end{array}$$

Second, derive $[\ : r \in r]$:

reject Restall's version of naive comprehension, but *so too* do the naive set theories explored in Restall (1992) and Priest (2006). The naive set theories of Brady (2006) and Weber (2012), on the other hand, reject Restall's extensionality principle, despite accepting their own versions. In other words: these going naive theories shy away from strong formulations of comprehension and extensionality. The present theory rejects only the rule of cut from Restall's list.

10 This can be shown by a strategy not unlike the one in footnote 5. Note in particular that no rule allows us to *instantiate* the quantifiers in a quantified sentence; no sequent rule removes any quantifiers that appear in its premise-sequents. The initial quantifiers in E, then, are 'stuck on' and cannot be put to use.

11 So-called because of the exciting uses to which it's put in Weber (2010, 2012). I should here be understood as introducing abbreviations; that is, where I write something like $[r \in r :\ ]$, that should be read as an abbreviation of $[\{x : x \notin x\} \in \{x : x \notin x\} :\ ]$. Recall that I'm using sets of formulas for my sequents; this means there will sometimes be contractions implicit in my derivations. For approaches to naive set theory that focus on blocking these contractions, see eg Grišin (1982), Petersen (2000), and Restall (1994).

$$\text{CompR: }\dfrac{\text{¬R: }\dfrac{[r \in r : r \in r]}{[\ \ : r \in r, r \notin r]}}{[\ \ : r \in r]}$$

The main derivation of [:] continues from these starting points, and can be found in Figure 2. So E2 is no use. (Note that the full derivation goes through in CL=Comp until the penultimate step. That much of it will be back later.)

3.3. The extensionality rule. Again, the key is to avoid a sentential formulation of extensionality altogether. The target strength of extensionality is given by the following rule, taken from Restall (2013):

$$\text{Ext: }\dfrac{[\Gamma, a \in t : a \in u, \Delta] \qquad [\Gamma, a \in u : a \in t, \Delta]}{[\Gamma : t = u, \Delta]}$$

As before, t and u here are any terms, while a is an eigenvariable. This allows for derivation of E as a theorem, as in Figure 3.

Unlike the attempt to add E as an axiom, though, Ext allows for derivation of $[\ \ : (\forall x(Px \equiv Qx)) \supset \{x : Px\} = \{x : Qx\}]$; the derivation is similar to the one in Figure 3, with a few applications of Comp rules inserted. Unlike the drop-rule approach, it remains nontrivial.[12]

3.3.1. Impure set theory. This is almost the final formulation. There is just one more feature that needs to be added, though; we must restrict the extensionality rule to *sets*. After all, extensionality is not a general principle of identity. Suppose Γ tells us that t and u

$$\text{⊥-drop: }\dfrac{\text{E-drop: }\dfrac{\text{∀L (x2): }\dfrac{\text{⊃L: }\dfrac{\text{CompR: }\dfrac{\text{=L2: }\dfrac{\text{CompR: }\dfrac{[\ \ : r \in r]}{[\ \ : z \in w]}}{[w = h : z \in h]}}{[w = h : \bot]} \qquad \text{∀R: }\dfrac{\text{∧R: }\dfrac{\text{⊃R: }\dfrac{\text{CompL: }\dfrac{\text{KR: }\dfrac{\text{⊥-drop: }\dfrac{[\bot : \bot]}{[\bot :\ \ ]}}{[\bot : z \in w]}}{[z \in h : z \in w]}}{[\ \ : z \in h \supset z \in w]} \qquad \text{⊃R: }\dfrac{\text{KR: }\dfrac{\text{CompL: }\dfrac{[r \in r :\ \ ]}{[z \in w :\ \ ]}}{[z \in w : z \in h]}}{[\ \ : z \in w \supset z \in h]}}{[\ \ : z \in w \equiv z \in h]}}{[\ \ : \forall z(z \in w \equiv z \in h)]}}{[\forall z(z \in w \equiv z \in h) \supset w = h : \bot]}}{[\forall x \forall y(\forall z(z \in x \equiv z \in y) \supset x = y) : \bot]}}{[\ \ : \bot]}}{[\ \ :\ \ ]}$$

Fig. 2. Deriving the empty sequent with E-drop.

$$\text{∀L (x2): }\dfrac{\text{⊃R: }\dfrac{\text{Ext: }\dfrac{\text{∀L: }\dfrac{\text{∧L: }\dfrac{\text{⊃L: }\dfrac{[a \in x : a \in x] \qquad [a \in y : a \in y]}{[a \in x \supset a \in y, a \in x : a \in y]}}{[a \in x \equiv a \in y, a \in x : a \in y]}}{[\forall z(z \in x \equiv z \in y), a \in x : a \in y]} \qquad \text{∀L: }\dfrac{\text{∧L: }\dfrac{\text{⊃L: }\dfrac{[a \in y : a \in y] \qquad [a \in x : a \in x]}{[a \in x \supset a \in y, a \in y : a \in x]}}{[a \in x \equiv a \in y, a \in y : a \in x]}}{[\forall z(z \in x \equiv z \in y), a \in y : a \in x]}}{[\forall z(z \in x \equiv z \in y) : x = y]}}{[\ \ : \forall z(z \in x \equiv z \in y) \supset x = y]}}{[\ \ : \forall x \forall y(\forall z(z \in x \equiv z \in y) \supset x = y)]}$$

Fig. 3. Deriving E.

[12] I won't quite show this here, as I'm about to make one final modification to the system, but the strategy I take in §4.2 could easily be adapted.

are both elephants, and further tells us that nothing is a member of any elephant. Then we will indeed have the premises of our extensionality rule. But we do not want to conclude that any two elephants are identical! It is only for sets that membership determines identity.

In a pure set theory, of course, there is no need to take account of this restriction; everything talked about in such a theory is a set. But a naive set theory should not *have* to be a pure set theory. It should work as an impure set theory as well. One of the main goals of a naive set theory is to allow for use of sets in an intuitive way to reason about any topic at all. For example, consider model-theoretic semantics for natural languages: it would be nice to be able to specify an *intended model*. The domain of such a model should be a set that includes everything. In some sophisticated set theories like ZF, there is no such set. In naive set theories, there can be, but only if the naive set theory is an impure one.[13]

One way to handle this is by including in our language a distinguished predicate S for '...is a set', and this is the route I'll pursue. With this predicate in hand, here are two possibilities for a restricted extensionality rule:

$$\text{ExtS1:}\quad \frac{[\Gamma, a \in t : a \in u, \Delta] \qquad [\Gamma, a \in u : a \in t, \Delta] \qquad [\Gamma : St, \Delta] \quad [\Gamma : Su, \Delta]}{[\Gamma : t = u, \Delta]}$$

$$\text{ExtS2:}\quad \frac{[\Gamma, a \in t : a \in u, \Delta] \qquad [\Gamma, a \in u : a \in t, \Delta]}{[\Gamma, St, Su : t = u, \Delta]}$$

ExtS2 follows from ExtS1, since we have $[St : St]$ and $[Su : Su]$ as axioms, and can weaken in any other needed side premises or conclusions. But the converse does not hold without cut. I don't see any particular reason to prefer either of these to the other, so I will work with ExtS1. Since this is the stronger rule, the nontriviality result below will cover both of them.

We also need to make sure, now that S is on the scene, that all set abstracts name sets. To do this, I'll use another drop rule:

$$S\text{-drop:}\quad \frac{[\Gamma, S\{x : A\} : \Delta]}{[\Gamma : \Delta]}$$

Finally, the full system is on stage: CL=Comp, plus ExtS1, plus S-drop. This is the target set theory. Call it NST, for 'naive set theory'.

§4. Models. In this section, I'll sketch a model-theoretic way to think about NST. Despite the affinities between the sequent formulation given above and classical logic, the model theory I'll present here is more closely tied to strong kleene logic and the paraconsistent logic LP. NST has its proof-theoretic foot in the classical and its model-theoretic foot in the nonclassical. (The same is true of the nontransitive approaches to other paradoxes I've cited above; this is what gives these approaches their of-both-worlds flavour.)

The models I'll use are three-valued models on the strong kleene scheme. That is, a model is a pair $\langle D, I\rangle$ of a domain and an interpretation, and I assigns members of the domain to terms, appropriate constructs of the usual sorts to predicates and relations, and values from the set $\left\{1, \frac{1}{2}, 0\right\}$ to sentences so that:

[13] Of course, there is more to this than can be said here; these remarks are purely motivational.

- $I(Pa_1a_2\dots a_n) = I(P)(\langle I(a_1), I(a_2), \dots, I(a_n)\rangle)$,
- $I(\bot) = 0$,
- $I(t = u) = 1$ iff $I(t) = I(u)$,
- $I(t = u) = I(u = t)$,
- $I(\neg A) = 1 - I(A)$,
- $I(A \wedge B) = \min(I(A), I(B))$,
- $I(A \vee B) = \max(I(A), I(B))$,
- $I(A \supset B) = \max(1 - I(A), I(B))$,
- $I(\forall x A(x)) = \min(I_x(A(x))$, for all x-variants I_x of I, and
- $I(\exists x A(x)) = \max(I_x(A(x))$, for all x-variants I_x of I.

To deal with NST, we need to require in addition that:

- $I(t \in \{x : A(x)\}) = I(A(t))$,
- $I(S\{x : A(x)\}) = 1$, and
- If $I(St) > 0$, $I(Su) > 0$, and $I(t = u) = 0$, then there is some $d \in D$ such that either $I(\in)(\langle d, I(t)\rangle) = 1$ and $I(\in)(\langle d, I(u)\rangle) = 0$ or else $I(\in)(\langle d, I(t)\rangle) = 0$ and $I(\in)(\langle d, I(u)\rangle) = 1$.

Each of these restrictions deals with one of the rules specific to NST. The first handles comprehension, the second handles S-drop, and the third handles ExtS1. Call a structure that meets all the other constraints, but possibly not these three, a *premodel*.

A (pre)model $\langle D, I\rangle$ *satisfies* a sequent $[\Gamma : \Delta]$ unless $I(\gamma) = 1$ for every $\gamma \in \Gamma$ and $I(\delta) = 0$ for every $\delta \in \Delta$, in which case it does not satisfy the sequent. A sequent is valid iff every model satisfies it. Note that, in order to fail to satisfy a sequent, a model must go from value 1 for the premises all the way to value 0 for the conclusions; a mere decrease from 1 to $\frac{1}{2}$ or from $\frac{1}{2}$ to 0 is not enough. As such, this definition of consequence is a bit out of the ordinary for polyvalent logics: it involves neither a notion of designated value nor a simple order-theoretic approach.

It's easy to show that NST is sound for these models: given the definitions above, every axiom is valid, and all rules preserve validity. I make no claim to completeness; when it matters, I'll distinguish NST proper from NSTm, where NSTm is the logic (whatever it is) that is determined by the above model theory. I'll show nontriviality directly for NSTm; it follows by soundness that NST too is nontrivial.

4.1. Comparison with LP. One more familiar approach to naive set theory is based on the logic LP, and is discussed in Restall (1992) and Priest (2006). This approach uses strong kleene models as well, but defines consequence differently, using the idea of designated values. Both $\frac{1}{2}$ and 1 count as designated values on an LP approach. Then a model LP-satisfies a sequent unless it assigns a designated value to every premise and an undesignated value (so 0) to every conclusion, and a sequent is LP-valid iff it's LP-satisfied by every model.

Despite this difference, the resulting set theory is very like NSTm in some regards. Crucially, the difference in the notion of satisfaction is only on the premise side; as a result, sequents with empty premises are either valid on both notions of consequence or neither. That is, the theorem-only fragments of the set theories are not affected by the difference in the way validity is defined. It is only in the applications to arguments with premises that the difference is revealed. It's long been known that LP behaves very classically in its theorem-only fragment, revealing its weakness only when considering arguments with premises. The approach I present here simply takes this classical aspect of an LP approach and extends it to the full logic.

This is just one manifestation of what is in fact a broad and tight connection between this type of nontransitive system on the one hand, and LP on the other. In addition to sharing models, these kinds of systems can also be captured simultaneously in a single proof-theoretic approach, as is explored in Ripley (2012), Hjortland (2013), and Barrio *et al.* (2014). In particular, Barrio *et al.* (2014) claim that this tight connection shows that there is not much benefit after all in any shift from LP to the kind of nontransitive systems that I recommend and have explored here. I will have to leave this issue for another day.

Differences in comprehension. There are some further slight differences between this approach and the particular ones given in Restall (1992) and Priest (2006), however. These are to do with the precise formulation of comprehension, and details around self-identity claims. The version of comprehension I've used here guarantees full intersubstitutivity between $A(t)$ and $t \in \{x : A(x)\}$ (in both NST and NSTm), but the version of comprehension used by Restall and Priest does not. Proof-theoretically, their version of comprehension amounts to using axioms $[A(t) : t \in \{x : A(x)\}]$ and $[t \in \{x : A(x)\} : A(t)]$ instead of the rules CompL and CompR; model-theoretically, it amounts to the requirement that $|I(A(t)) - I(t \in \{x : A(x)\})| < 1$, rather than the stricter requirement I've imposed.

As I mentioned in footnote 8, the stricter requirement I impose here can be seen as a reflex of a certain kind of deflationism about sets, patterned on the kinds of approaches taken to truth in Beall (2009) and Field (2008). (I don't claim, of course, that this is the *only* way to motivate the stricter requirement.) I set discussion of the relative ups and downs of these approaches aside here.[14]

Differences in extensionality. Restall also defines identity out of $\in$, rather than building it into the logic itself, as I have here; and both Restall and Priest consider only pure set theories. But to the extent that they are comparable, the approach to identity enshrined in NST and NSTm is again stronger. The LP-based approaches Restall and Priest consider require only that $I(t = t) \geq \frac{1}{2}$, rather than that it be 1, as NSTm does. This more or less amounts to the proof-theoretic difference between NST's $=$-drop rule and its weaker axiom relative.[15]

An anonymous referee suggests two other possible treatments of extensionality: first, adding axiomatic sequents of the form $[\forall z(z \in t \equiv z \in u) : t = u]$; or second, adding the following rules:

$$\frac{[\Gamma, \forall z(z \in t \equiv z \in u) : \Delta]}{[\Gamma, t = u : \Delta]} \qquad \frac{[\Gamma : \forall z(z \in t \equiv z \in u), \Delta]}{[\Gamma : t = u, \Delta]}$$

14 It is worth noting, however, that the more stringent form of comprehension I adopt here has its costs as well. Foremost among these is that the only LP model meeting the strong comprehension condition that is also a model of ZF is the trivial model: the one assigning value $\frac{1}{2}$ to every formula. (This has been proved by Morgan Thomas, who called the fact to my attention.) So the strong comprehension condition blocks a certain kind of recapture result used, for example, in Priest (2006, chap. 18). It is also incompatible with the extensions of LP considered in Omori (2014).

15 $=$-drop does not on its own guarantee that $I(t = t) = 1$ on every model, but it gets close: it guarantees that if there is a countermodel at all to an argument, then there is a countermodel $\langle D, I\rangle$ to that argument such that $I(t = t) = 1$ for every t. This means that we can safely ignore any models on which $I(t = t) \neq 1$, as they will not affect the resulting consequence relation.

(I ignore the extension to impure set theory here; either suggestion could easily be adapted to such a setting.) While the issue deserves more discussion than I can give it here, I suspect that in the end neither suggested route would prove very attractive, largely owing to the presence of $\forall$ and $\equiv$. Without a rule of cut, it is very difficult to exploit information contained in complex formulas. To some extent, this is the reason for the weakness of the axiomatic formulations of comprehension and extensionality considered earlier. I suspect that these suggestions would run into similar difficulties, although the suggestions certainly deserve exploration. It may also be worth noting that the axiomatic sequents envisioned in the first suggestion are already provable from the formulation of extensionality I've adopted here.

So there are some fiddly differences that require care; presentations of LP-based set theories have tended to use slightly weaker principles of comprehension and self-identity than the ones I've deployed in this paper. But there is no need for them to have adopted these weaker requirements in order to avoid triviality, as I'll show presently. The nontriviality proof for NSTm in §4.2 is equally a nontriviality proof for an LP-style reformulation, with the stronger versions of these principles. The proof works by arriving at a model, and the model is just as much a model of the LP-style reformulation as it is of NST itself. While I happen to think that NST gives a superior way to reason about and with sets, it's certainly open to the LP partisan to help themselves to this nontriviality proof as well, and thereby achieve tighter formulations of comprehension and self-identity than they've had in the past; indeed, I hope they do just that.

4.2. Nontriviality. In this section, I'll show that NSTm is nontrivial. Since NST is sound with respect to NSTm, it too must be nontrivial. In fact, I'll show something stronger than nontriviality: quantifier-free model-theoretic conservativity. Here's what I mean by that: start from a strong kleene model M for the fragment of our language that excludes $\in$, S, and the set-abstract terms $\{x : A(x)\}$. (I'll call this fragment the *non-set vocabulary*.) The proof below will show how to expand that model to a model M' for the full vocabulary, with the following feature: for any quantifier-free sentence A of the non-set vocabulary, the value A receives in M' is the same as the value it receives in M. (The transition from M to M' will add a bunch of things to the domain, to play the role of sets; this may mess with quantified sentences in the non-set vocabulary, but it won't affect quantifier-free sentences.)

From this, nontriviality quickly follows. Moreover, this doesn't just show that there is *some* argument that's not valid: it shows that a very wide range of arguments are not valid. Indeed, any quantifier-free argument in the non-set vocabulary that is not already valid in classical logic is also not valid in NSTm, and so not in NST.[16]

The proof will follow the strategy of Brady (1971) and similar approaches for comprehension, but will deal with extensionality in a new (and simpler) way. It works by defining a starting point and a jump operation, and then showing that, if the jump operation is iterated into the transfinite, it will eventually reach a fixed point: a point at which further jumps just stay still. The starting point is determined quickly from M, and the eventual fixed point will be M'. Let's begin!

[16] This follows from what goes below, plus the fact that CL= is not only sound for strong kleene models of the non-set vocabulary, but also complete. For proof of this, see Ripley (2012).

4.2.1. Starting point. Begin with a model $M = \langle D, I \rangle$ of the non-set vocabulary that meets all the requirements in §§4 (except, of course, for the requirements specific to the set vocabulary). I will assume that M contains a term for every member of its domain. The first step is to add things to the domain to play the roles of sets. If the set abstract terms themselves are not already in the domain, they will do fine; this is particularly convenient, so I'll go with it. (If the abstract terms *are* already in the domain, then just swap them out for something else first.) Let Abs be $\{t : t \text{ is a set abstract term}\}$. (Note that Abs is a (informal ZFC) set of object-language terms, not itself a set abstract term of the object language.)

Let $M_0 = \langle D', I_0 \rangle$, where D' is $D \cup Abs$, and I_0 extends I to the full language as follows:

- $I_0(S) = \{\langle t, 1 \rangle : t \in Abs\} \cup \{\langle t, 0 \rangle : t \notin Abs\}$,
- For $t \in Abs$, $I_0(t) = t$,
- For $t \in Abs, u \notin Abs$, $I_0(=)(\langle t, u \rangle) = 0$,
- For distinct $t, u \in Abs$, $I_0(=)(\langle t, u \rangle) = \frac{1}{2}$,
- For $t \notin Abs$, $I_0(\in)(\langle u, t \rangle) = 0$,
- For $t \in Abs$, $I_0(\in)(\langle u, t \rangle) = \frac{1}{2}$,
- For $t \in Abs$ and R an $n$ary non-set relation other than $=$, $I_0(R)(\langle a_1, \dots, t, \dots a_n \rangle) = \frac{1}{2}$.

In words, we ensure that S has the right extension; that each set abstract term names itself; that $=$ takes value 0 on claims of identity between sets and nonsets, and $\frac{1}{2}$ on claims of identity between sets named by distinct terms; that $\in$ takes value 0 when its right relatum is not a set abstract and $\frac{1}{2}$ when it is; and that non-set predicates other than $=$ always take value $\frac{1}{2}$ on set abstracts. (This last may seem odd, but recall that this is merely a nontriviality proof; we are not in any way constructing an intended model here.)

M_0 is a premodel; it may well not be (is probably not!) a model of the full vocabulary, as it has taken no special precautions to meet the constraints for comprehension or ExtS1. (It has taken care of S-drop, though.) This is the reason for the construction that follows.

4.2.2. The construction. From M_0, there is no further need to change the domain or the extension of S; these are already as they will be in M'. Nor is there any need to modify the extension of any non-set predicate other than $=$; these too will stay put. All that will need to be tweaked is the extension of the predicates $\in$ and $=$. Even these, though, will only need to change in certain ways. Nothing needs to be a member of a non-set, so we only need to worry about cases of $\in$ where the right relatum is a set abstract. And no set needs to be identical to any non-set, so we only need to worry about cases of $=$ where both relata are set abstracts.

We change these extensions via a construction that builds a transfinite sequence of premodels. For successors $n+1$, $M_{n+1} = \langle D', I_{n+1} \rangle$, where I_{n+1} is related to I_n as follows (note that D' stays untouched):

- For all terms t, $I_{n+1}(t) = I_n(t)$,
- For all predicates and relations R other than $\in$ and $=$, $I_{n+1}(R) = I_n(R)$,
- For $t \notin Abs$, $I_{n+1}(\in)(\langle d, t \rangle) = I_n(\in)(\langle d, t \rangle)$,
- For $t \notin Abs$, $I_{n+1}(=)(\langle d, t \rangle) = I_n(=)(\langle d, t \rangle)$,
- For $t, u \in Abs$, if there is a $d \in D$ such that $I_n(\in)(\langle d, t \rangle) = 1$ and $I_n(\in)(\langle d, u \rangle) = 0$, then $I_{n+1}(=)(\langle t, u \rangle) = 0$; otherwise, $I_{n+1}(=)(\langle t, u \rangle) = I_n(=)(\langle t, u \rangle)$,

- For $t \in Abs$ such that t is the term $\{x : A(x)\}$, $I_{n+1}(\in)(\langle I_{n+1}(u), t\rangle) = I_n(A(u))$.[17]

As is quick to verify, these clauses specify a unique premodel for each successor stage. It remains only to specify the limit stages. We will need a tiny bit more machinery: the so-called *information order* $\sqsubseteq$ on our three values. This is the partial order $\sqsubseteq$ such that $\frac{1}{2} \sqsubseteq 1$; $\frac{1}{2} \sqsubseteq 0$; and 0 and 1 are $\sqsubseteq$-incomparable. Where $V \subseteq \{0, \frac{1}{2}, 1\}$, let $\max_{\sqsubseteq}(V)$ be the $\sqsubseteq$-maximum value in V; note that this is well-defined only when $\{1, 0\} \not\subseteq V$.

Where n is a limit ordinal, $M_n = \langle D', I_n\rangle$, where I_n is related to its predecessors as follows:

- For all terms t, $I_n(t) = I_m(t)$, for any $m < n$,
- For all predicates and relations R other than $\in$ and $=$, $I_n(R) = I_m(R)$, for any $m < n$,
- For $t \notin Abs$, $I_n(\in)(\langle d, t\rangle) = I_m(\in)(\langle d, t\rangle)$, for any $m < n$,
- For $t \notin Abs$, $I_n(=)(\langle d, t\rangle) = I_m(=)(\langle d, t\rangle)$, for any $m < n$,
- For $t, u \in Abs$, $I_n(=)(\langle t, u\rangle) = \max_{\sqsubseteq}(\{I_i(=)(\langle t, u\rangle)\}_{i<n})$,
- For $t \in Abs$, $I_n(\in)(\langle u, t\rangle) = \max_{\sqsubseteq}(\{I_i(\in)(\langle u, t\rangle)\}_{i<n})$.

The first four bullet points here are easy: these are the values that remain constant through the entire construction, so we simply hold them constant at limit stages as well. This constancy is required for I_n to be well-defined here, but it is obvious, since nothing in the construction can change these values. The second two bullet points are a bit trickier: they are only well-defined if the I_is behave properly for i below n. In fact they do, but this requires proof.

4.2.3. Increasingness and a fixed point. This proof is wrapped up in the proof of a more general fact: *increasingness*.

LEMMA 4.1 (Increasingness). *For any sentence A and any ordinals m, n such that $m \leq n$, $I_m(A) \sqsubseteq I_n(A)$.*

Proof. Proof is by induction on n. When $m = n$ (including when $n = 0$), we're all set; so assume $m < n$. Now we have two inductive steps: one for n a successor, and the other for n a limit. Each of these steps itself proceeds by induction on A's formation. Here the inductive steps (eg $\wedge, \neg, \forall$) are straightforward and the base cases (for atomic A) are where all the action is. As such, I'll skip the inductive steps and focus on the base cases.

When A is atomic, there are four cases, depending on A: 1) A is $u \in t$, for $t \in Abs$, 2) A is $t = u$, for distinct $t, u \in Abs$, 3) A is $t = t$, for $t \in Abs$, or 4) A is some other atomic formula. In the third case, we're all set; no formula of the form $t = t$ ever takes a value other than 1 in the course of the construction (since the stages are all premodels). In the fourth case, we're also all set: no atomic formula of type 4 changes its value anywhere in the construction, as is easy to spot from §4.2.2. So we only need to worry about the first two cases. In fact, we only need to distinguish between these cases when n is a successor; for limit n the same reasoning works for both.

Suppose, then, that n is a successor $k + 1$. We know $m \leq k$, so by the (main) inductive hypothesis, $I_m(A) \sqsubseteq I_k(A)$. It remains only to show, then, that $I_k(A) \sqsubseteq I_n(A)$.

[17] This is where it becomes important that M contained a term for every member of its domain; without such an assumption, this clause would not fully characterize $I_{n+1}(\in)$ (although a more complicated clause could be rigged up to do the trick).

If $I_k(A) = \frac{1}{2}$, we're done. So suppose $I_k(A) = 1$ or 0. Now we case on A, and show that $I_n(A) = I_k(A)$:

Case 1. A is $u \in t$, where t is the term $\{x : B(x)\}$. Since $I_k \neq \frac{1}{2}$, it follows that there is an $i < k$ such that $I_k(A) = I_i(B(u))$. By the inductive hypothesis, $I_k(B(u)) = I_i(B(u))$, and by the construction, $I_n(A) = I_k(B(u))$, so $I_n(A) = I_k(A)$.

Case 2. A is $t = u$, for distinct $t, u \in Abs$. There is nothing in the construction that could allow $I_k(A)$ to be 1, so $I_k(A)$ must be 0. This can only happen, though, when there is some $i < k$ and some $d \in D'$ such that $I_i(\in)(\langle d, t\rangle) = 1$ and $I_i(\in)(\langle d, u\rangle) = 0$, or vice versa. So by the inductive hypothesis, $I_k(\in)(\langle d, t\rangle) = 1$ and $I_k(\in)(\langle d, u\rangle) = 0$, or vice versa. But then by the construction, $I_n(=)(\langle t, u\rangle) = 0$, and so $I_n(A) = 0$.

So much for the successor case. Now suppose n is a limit ordinal, and that either A is $u \in t$, for $t \in Abs$, or else A is $t = u$, for $t, u \in Abs$. By the inductive hypothesis, there cannot be $j \leq k < n$ such that $I_j(A) \not\sqsubseteq I_k(A)$. As such, $\max_{\sqsubseteq}(\{I_i(A)\}_{i<n})$ is well-defined, so by the construction $I_n(A) = \max_{\sqsubseteq}(\{I_i(A)\}_{i<n})$. But $I_m(A) \sqsubseteq \max_{\sqsubseteq}(\{I_i(A)\}_{i<n})$, so $I_m(A) \sqsubseteq I_n(A)$. □

So as we climb the transfinite construction, formulas might move from value $\frac{1}{2}$ to either value 1 or value 0, but once a formula is at either 1 or 0, it never moves again, no matter how high we climb. Given this, and the fact that there are only so many formulas in the language, we have:

COROLLARY 4.2 (Fixed point). *There is some ordinal n such that* $M_n = M_{n+1}$.

Proof. If for every n, $M_n \neq M_{n+1}$, then, since the M_ns differ only in their I component, it follows that for every n there is some formula A such that $I_n(A) \neq I_{n+1}(A)$. Moreover, this must be a distinct A for each n; since $I_n(A) \sqsubseteq I_{n+1}(A)$, $I_{n+1}(A)$ must be 1 or 0, and so $I_m(A)$ must equal $I_{n+1}(A)$ for all $m \geq n+1$. But there are not enough formulas to have a distinct A for every ordinal n. Contradiction. □

4.2.4. Nontriviality. Let $M' = \langle D', I'\rangle$ be the fixed point reached by the construction. That is, $M' = M_n$ for some n such that $M_n = M_{n+1}$. It remains to be shown that M' is not merely a premodel, but is a model as well.

LEMMA 4.3. *M' is a model.*

Proof. There are three conditions to check.

- First, that $I'(t \in \{x : A(x)\}) = I'(A(t))$. We know that $I_{n+1}(t \in \{x : A(x)\}) = I_n(A(t))$, for any n. Then it is immediate that the condition is met, since there is some n such that $I' = I_n = I_{n+1}$.
- Second, that $I'(S(\{x : A(x)\})) = 1$. We know that $I_0(S(\{x : A(x)\})) = 1$, and then it follows by increasingness that the condition is met.
- Third, that if $I'(St) > 0$, $I'(Su) > 0$, and $I'(t = u) = 0$, then there is some $d \in D'$ such that either $I'(\in)(\langle d, I'(t)\rangle) = 1$ and $I'(\in)(\langle d, I'(u)\rangle) = 0$, or else $I'(\in)(\langle d, I'(t)\rangle) = 0$ and $I'(\in)(\langle d, I'(u)\rangle) = 1$.

  We know that $I_0(Sv) = 0$ unless $v \in Abs$; it follows by increasingness that $I'(Sv) = 0$ unless $v \in Abs$. So we only need to worry about the case where $t, u \in Abs$. In this case, $I'(t) = t$ and $I'(u) = u$.

  If $I_{n+1}(t = u) = 0$, then t is distinct from u, and there is some $k < n+1$ and $d \in D'$ such that $I_k(\in)(\langle d, t\rangle) = 1$ and $I_k(\in)(\langle d, u\rangle) = 0$, or vice versa.

(Otherwise, $I_{n+1}(t = u)$ would still be $\frac{1}{2}$, as $I_0(t = u)$ is.) By increasingness, $I_n(\in)(\langle d, t\rangle) = 1$ and $I_n(\in)(\langle d, u\rangle) = 0$, or vice versa. But since there is some n such that $I' = I_n = I_{n+1}$, it follows that the condition is met. □

Moreover, for every quantifier-free A in the non-set vocabulary, $I'(A) = I(A)$, as is easy to check. So we have quantifier-free model-theoretic conservativity:

COROLLARY 4.4 (Quantifier-free conservativity). *Every model for the non-set language can be extended to a model for the full language such that the two models do not differ on the values they assign to any quantifier-free formula of the non-set language.*

From this, nontriviality is immediate:

COROLLARY 4.5 (Nontriviality). *NSTm and NST are nontrivial.*

Proof. It's easy to build a non-set countermodel M to the argument from Pa to Qb. But Pa and Qb have no quantifiers, so M can be extended to a model M' of the full language such that $M'(Pa) = M(Pa)$ and $M'(Qb) = M(Qb)$. Since M was a countermodel to the argument, so too will M' be. Thus, NSTm is nontrivial. Since NSTm is at least as strong as NST, NST too is nontrivial. □

4.3. Discussion. It's important to note that the model M' constructed in this nontriviality proof is not in any way supposed to play the role of an *intended extension* of M; it does not capture the actual structure of naive sets. (Nor should we expect it to, as it is itself a construction in a sophisticated set theory.) For example, in an M' produced by the above construction, the terms $\{x : Px \vee Qx\}$ and $\{x : Qx \vee Px\}$ will denote different members of the domain, since each denotes itself and they are, after all, distinct terms.

Presumably, though, the set of all x such that $Px \vee Qx$ is the same set as the set of all x such that $Qx \vee Px$. Indeed, NST allows for this identity to be derived. Don't think that the model produced by the nontriviality proof has more importance than it does! Its purpose is only to show that NST is nontrivial (if classical ZFC is). Similar comments apply to $\{x : A(x)\}$ and $\{y : A(y)\}$; although these terms denote distinct members of the domain of M', the sentence $\{x : A(x)\} = \{y : A(y)\}$ remains a theorem of NST.

However, theorem-strength identity is only so strong. It follows from the considerations in §3.2 that the claim that the weber set is identical to the empty set is also a theorem of NST; but still these two sets behave differently. (For example, it's a theorem that x is a member of the weber set; but it's definitely *not* a theorem that x is a member of the empty set!)

We might want a stronger connection between $\{x : Px \vee Qx\}$ and $\{x : Qx \vee Px\}$ than this. Fortunately, this can be achieved by strengthening NST slightly. Let R be some equivalence relation on set abstract terms such that $\{x : A(x)\}R\{y : B(y)\}$ only if $I(A(a)) = I(B(a))$ on every premodel.[18]

Given such an R, we can now strengthen NST to NST^R by adding the following rule, for set abstract terms t, u such that tRu:

$$R\text{-=-drop:}\quad \frac{[\Gamma, t = u : \Delta]}{[\Gamma : \Delta]}$$

[18] This condition amounts to equivalence between $A(a)$ and $B(a)$ in the logic sometimes called S3 and sometimes called first-degree RM.

NST^R can be proved nontrivial in the same way as NST, mutatis mutandis; in setting up the domain for the construction, rather than including the set abstract terms themselves, we just need to include their equivalence classes under R. The constraint on R ensures that the construction remains well-defined; the proof goes through just as before, mutatis mutandis.

For any R meeting the constraint, if tRu, then NST already derives $[\ : t = u]$. But recall that being able to drop a formula from the premises of an argument is a stronger condition to meet than simply being able to conclude the formula as a theorem. It amounts to concluding it as a theorem *and then cutting on it*. As such, NST^R imposes a stronger connection between R-related set abstract terms than NST does. Given the $= L$ rules, being able to drop $t = u$ yields full intersubstitutivity between t and u: a very strong connection indeed.

If we go down this path, I think one natural candidate for such an R is the strongest possible choice. On this choice, the constraint on R is strengthened to a biconditional: $\{x : A(x)\}R\{y : B(y)\}$ *if and* only if $I(A(a)) = I(B(a))$ on every premodel. This choice of R yields full intersubstitutivity between $\{x : Px \vee Qx\}$ and $\{x : Qx \vee Px\}$, as well as between $\{x : A(x)\}$ and $\{y : A(y)\}$, and between many more pairs besides.

Perhaps there is good reason to choose one R rather than another, or to avoid the issue altogether and stick with the original NST. For now, I'm happy just to put them on the table: they are all nontrivial.

An anonymous referee suggests here an additional, stronger, desideratum: "that, whenever the ground model for the non-set language verifies the coextensionality of $\phi(x)$, and $\psi(x)$, the final model assigns 1 to the identity $\{x : \phi(x)\} = \{x : \psi(x)\}$". I do not know whether this is achievable in the present setting, and it would indeed be nice to know! Even with NST^R in play, there is no connection this tight being enforced.[19]

§5. Conclusion. This paper has extended the nontransitive approach to paradoxes explored for vagueness in Cobreros *et al.* (2012) and Ripley (2013b) and for truth in Ripley (2013a), Ripley (2012), and Cobreros *et al.* (2013) to the paradoxes engendered by a naive set theory: a set theory characterized by the two principles of comprehension and extensionality. The key to such an extension is in the use of rule-based formulations of these principles, rather than sentence-based formulations.

The resulting set theory, in its theorem fragment, resembles the LP-based set theories explored in Restall (1992) and Priest (2006), but with a stronger form of comprehension, and stronger self-identity. The full approach, however, is classical-ish as well as LP-ish, as it validates every classically-valid argument.

Despite this, it remains nontrivial, as has been shown via a modification of the technique of Brady (1971). The main innovation in this proof is in the treatment of extensionality, which in the present context, unlike in many others, is no harder to work with than comprehension.

§6. Acknowledgments. For helpful discussion, thanks to Jc Beall, Ross Brady, Rohan French, Hitoshi Omori, Graham Priest, Greg Restall, Lionel Shapiro, Morgan Thomas,

[19] I believe, but am not certain, that the referee's desideratum can be achieved for restricted ϕ and ψ—say, without = or quantifiers—without too much trouble, by modifying the specification of M_0.

Ross Vandegrift, Zach Weber, the Melbourne Logic Group, and the UConn Logic Group. Two anonymous referees for this journal, and two for another, all also provided very helpful comments. This research was partially supported by the Government of Spain via the grant *Non-Transitive Logics*, number FFI2013-46451-P.

BIBLIOGRAPHY

Barrio, E., Rosenblatt, L., & Tajer, D. (2014). The logics of strict-tolerant logic. *Journal of Philosophical Logic*. To appear.

Beall, J. (2009). *Spandrels of Truth*. Oxford: Oxford University Press.

Brady, R. T. (1971). The consistency of the axioms of abstraction and extensionality in a three-valued logic. *Notre Dame Journal of Formal Logic*, **12**(4), 447–453.

Brady, R. T. (1989). The non-triviality of dialectical set theory. In Priest, G., Routley, R., and Norman, J., editors. *Paraconsistent Logic: Essays on the Inconsistent*. Münich: Philosophia Verlag, pp. 437–471.

Brady, R. T. (2006). *Universal Logic*. Stanford, California: CSLI Publications.

Cobreros, P., Égré, P., Ripley, D., & van Rooij, R. (2012). Tolerant, classical, strict. *Journal of Philosophical Logic*, **41**(2), 347–385.

Cobreros, P., Égré, P., Ripley, D., & van Rooij, R. (2013). Reaching transparent truth. *Mind*, **122**(488), 841–866.

Field, H. (2008). *Saving Truth from Paradox*. Oxford: University Press.

Field, H. (2014). Naive truth and restricted quantification: Saving truth a whole lot better. *Review of Symbolic Logic*, **7**(1), 147–191.

Field, H., Lederman, H., & Øgaard, T. F. (2014). Prospects for a naive theory of classes. To appear.

Glanzberg, M. (2005). Minimalism, deflationism, and paradoxes. In Beall, J. and Armour-Garb, B., editors. *Deflationism and Paradox*. Oxford: Oxford University Press, pp. 107–132.

Grišin, V. N. (1982). Predicate and set-theoretic calculi based on logic without contractions. *Mathematics of the USSR—Izvestiya*, **18**(1), 41–59. (English translation).

Hinnion, R., & Libert, T. (2003). Positive abstraction and extensionality. *Journal of Symbolic Logic*, **68**(3), 828–836.

Hjortland, O. (2013). Logical pluralism, meaning variance, and verbal disputes. *Australasian Journal of Philosophy*, **91**(2), 355–373.

Incurvati, L. (2012). How to be a minimalist about sets. *Philosophical Studies*, **159**(1), 69–87.

Negri, S., & von Plato, J. (1998). Cut elimination in the presence of axioms. *The Bulletin of Symbolic Logic*, **4**(4), 418–435.

Omori, H. (2014). Remarks on naive set theory based on LP. *Review of Symbolic Logic*. To appear.

Petersen, U. (2000). Logic without contraction as based on inclusion and unrestricted abstraction. *Studia Logica*, **64**(3), 365–403.

Priest, G. (2006). *In Contradiction*. Oxford: Oxford University Press.

Restall, G. (1992). A note on naïve set theory in LP. *Notre Dame Journal of Formal Logic*, **33**(3), 422–432.

Restall, G. (1994). *On Logics Without Contraction*. PhD thesis, The University of Queensland.

Restall, G. (2013). Assertion, denial, and non-classical theories. In Tanaka, K., Berto, F., Mares, E., and Paoli, F., editors. *Paraconsistency: Logic and Applications*. Dordrecht: Springer, pp. 81–100 .

Ripley, D. (2012). Conservatively extending classical logic with transparent truth. *Review of Symbolic Logic*, **5**(2), 354–378.

Ripley, D. (2013a). Paradoxes and failures of cut. *Australasian Journal of Philosophy*, **91**(1), 139–164.

Ripley, D. (2013b). Revising up: Strengthening classical logic in the face of paradox. *Philosophers' Imprint*, **13**(5), 1–13.

Routley, R. (1977). Ultralogic as universal? *Relevance Logic Newsletter*, **2**, 51–89. Reprinted in Routley (1980).

Shapiro, L. (2011). Deflating logical consequence. *Philosophical Quarterly*, **61**(243), 320–342.

Weber, Z. (2010). Transfinite numbers in paraconsistent set theory. *Review of Symbolic Logic*, **3**(1), 71–92.

Weber, Z. (2012). Transfinite cardinals in paraconsistent set theory. *Review of Symbolic Logic*, **5**(2), 269–293.

White, R. B. (1979). The consistency of the axiom of comprehension in the infinite-valued predicate logic of Łukasiewicz. *Journal of Philosophical Logic*, **8**(1), 509–534.

DEPARTMENT OF PHILOSOPHY
101 MANCHESTER HALL
344 MANSFIELD RD
UNIVERSITY OF CONNECTICUT
STORRS, CT 06269 USA
E-mail: davewripley@gmail.com