SETS & SUPERSETS

TOBY MEADOWS

You're gonna need a bigger boat.
(Jaws)

ABSTRACT. It is a commonplace of set theory to say that there is no set of all well-orderings nor a set of all sets. We are implored to accept this due to the threat of paradox and the ensuing descent into unintelligibility. In the absence of promising alternatives, we tend to take up a conservative stance and tow the line: there is no universe [Halmos, 1960]. In this paper, I am going to challenge this claim by taking seriously the idea that we *can* talk about the collection of all the sets and many more collections beyond that. A method of articulating this idea is offered through an indefinitely extending hierarchy of set theories. It is argued that this approach provides a natural extension to ordinary set theory and leaves ordinary mathematical practice untouched.

The idea that there ought to be a set containing all of the sets is the focus of this paper. It is not a new idea. Moreover, I suspect that my way of addressing this issue will have occurred to many readers, although usually they they will have set it aside. I do not want to suggest that the techniques explored in this paper are of themselves particularly original; we shall see that related approaches have been explored from a technical angle for many years. Rather, what I am aiming to do is give this group of ideas a more philosophical twist by drawing out their underlying motivations and taking the resultant view to its natural conclusion. Beyond this, I will also investigate why such approaches are so often rejected and argue that we may adopt a more sympathetic attitude toward them.

The first section of the paper begins by examining the motivations that could lead someone to accept that such sets exist and finishes with an exposition of my proposal for making sense of this. The implementation will be to extend rather than revise the orthodox set theoretic foundation of ZFC. In the second section, we examine alternative solutions to the problem and show that this contest has no clear winner. Finally, I shall then defend the proposal from some objections and problems. In particular, I respond to what I take to be the most serious objection: a misplaced deference to mathematical expertise. I shall argue that while the proposal should be interesting to philosophers, mathematicians may have good reason to ignore it.

I would like to thank Zach Weber, Øystein Linnebo, Carrie Jenkins, James Studd, Stephen Read, Volker Halbach, Jc Beall, Dan Isaacson, Torfinn Huvenes, Kentaro Fujimoto for providing invaluable assistance in the development of this paper. I would also like to thank Oxford University and the University of St Andrews for giving me the opportunity to present these ideas. Finally, I would like to thank two anonymous referees for their incisive comments and suggestions for the paper.

1. Motivating supersets

1.1. **A Problem.** Our underlying motivation stems from a problem emerging with the notion of an indefinitely extensible concept. Dummett provides us with the canonical formulation:

> *[an] indefinitely extensible concept is one such that, if we can form a definite conception of a totality all of whose members fall under the concept, we can, by reference to that totality, characterise a larger totality all of whose members fall under it.*[1] *[Dummett, 1993]*

To make this clearer, it will be helpful to consider a couple of plausible examples of indefinitely extensible concepts:

(1) natural numbers; and
(2) sets.

With the first example, we shall catch a glimpse of the motivating problem, but in a setting in which it is easy to *solve*. In the second case, the underlying problem of the paper will emerge.

1.1.1. *Natural numbers.* The ordinary process of counting should remind us of indefinite extensibility. Suppose X is a *definite collection* of, say, sheep. For the sake of our illustration, let us say that it is definite in the sense of being one that we have just counted. We can then form a larger collection by counting the first sheep we see that is not in our original *totality*. Thus we get a *larger totality* which additionally includes a new sheep. A similar process may be used to count out the natural numbers themselves. Given the totality of the first, say five, natural numbers we might form the larger totality of the first six natural numbers by adding the least number not already included in the original totality. Thus, there is a sense in which the natural numbers are indefinitely extensible.

But this leads us to a kind of problem. What are these natural numbers and how can we refer to their totality? The kind of reference envisaged here is *collective* rather than *distributive*. We are interested in talking about the totality itself of natural numbers, not merely that members of that totality.[2] Now given that we are assuming that an collection becomes definite only when it can be counted, the natural numbers cannot be definite. Without being a definite totality, it is difficult to see how one might refer to the natural numbers given that it appears to have no fixed extension.[3] The cause of this problem is the fact that the natural numbers are indefinitely extensible.

Of course, it is possible for us to refer to the totality of natural numbers. I just did. For this we have Cantor's paradise of sets to thank. We collect together all of the natural numbers into a *set*. In this way, we obtain a definite totality rather than an indefinite one.

There is, however, a certain sleight of hand here: the meaning of *definite* has been allowed to shift.[4] We originally supposed that a collection was definite if we - at least in principle - could finish counting it. But in the world of sets we are able to fix on a collection as long as we can find sufficient means for picking it out.

---

[1] We should note that the formulation has some problems as discussed in Shapiro and Wright's [2007], however, these issues can be largely ignored for the purposes of this paper.

[2] Indeed all discussion of reference to collections in this paper will be collective rather than distributive unless stated otherwise. It should also be noted that we shall sometimes use reference to pluralities in place of collective reference. For example, in speaking about the totality itself of natural numbers, we speak of referring to *them* rather than *it*. For our purposes, we take it that articulations of either kind are interchangeable, although we understand that this assumption is controversial (see Section 2.2).

[3] For further discussion of this and related points see [Dummett, 1993: p439].

[4] For a detailed discussion of relativised notions of definiteness and their impact on indefinite extensibility see [Shapiro and Wright, 2007].

Given that we do appear to be able to refer to the collection of natural numbers, we may have reason to think that linking definiteness with being able to be counted was an error.

1.1.2. *Sets.* In opening ourselves out onto the world of sets and infinite collections, problems like the previous one become trivial. Indeed, we quickly gain the ability to talk about and manipulate all of the objects investigated in mathematics. But in this context, a related problem of indefinite extensibility has more sting. We first show that the sets, like the natural numbers, can be considered to be indefinitely extensible; then we consider the analogue of the previous problem as it applies to sets.

Let x be an arbitrary set among whose members none is a member of itself.[5] This is would appear to be a *definite totality*. Then let us consider the set-theoretic property of *not being a member of itself*. Using this property and *by reference to this totality* we may form a *larger totality* as follows:

(1) Given our set x, take the set z of those members $y \in x$ such that $y \notin y$. Clearly, z is just x itself.
(2) Now form a new set comprised of the members of x and x itself; i.e., $x \cup \{x\}$.
(3) It is then clear that $x \cup \{x\}$ is also a set that contains no object which is a member of itself; thus, we may repeat the process.[6]

This tells us that the sets are indefinitely extensible.

Now let us consider an analogous problem to the one we had with the natural numbers. Just as we wanted to talk about the collection of all natural numbers, so too we may want to talk about the collection of all sets. For example, we might want to say that every set is contained in it. The natural thing to do is see if we can repeat or generalise Cantor's trick. The obvious approach appears disastrous.

Suppose we gather together all of the sets which are not members of themselves. Call this set R. Naïvely, it seems obvious that R is a set: it is just another collection. But then ask whether R is a member of itself? If we suppose it is, then by definition it is not a member of itself. Thus our assumption was wrong and we may conclude that it is not a member of itself. But then it satisfies the conditions for being a member of R. Thus it is both a member of itself and not a member of itself: contradiction. This is Russell's paradox.[7]

The traditional solution to this problem is to say that R was not a set after all; or in Dummett's terminology, we might say that R does not form a definite totality. For example, a set theory like ZFC takes up this approach by restricting its comprehension axiom. Rather than accepting that just any definition gives us a set, we only allow definitions to generate new sets by *separating out* the objects that satisfy the definition from a pre-existing set. Thus, we have the axiom of separation:

$$\forall w \exists x \forall z (z \in x \leftrightarrow z \in w \land \varphi(z)).$$

[5]This is a standard demand in a mainstream set theory like ZFC. In fact, the Axiom of Foundation prohibits there being any self-membered sets.

[6]If we had started this process with the empty set and continued as above, we would end up with the von Neumann ordinals.

[7]We can also see this problem emerging from the definition of indefinite extensibility. Let R be as above. Then R is a *definite totality*, so by *reference to that totality* we should be able to form a *larger totality* than R. Perhaps we try to form $R \cup \{R\}$, which should also be a set. But this cannot be right. We can see that $R \notin R$, so we must have $R \cup \{R\} \supsetneq R$; and yet R already contains all the non-self-membered sets: this is impossible.

The upshot of this is that while the problem for the natural numbers could be solved using sets, the analogous problem for the sets themselves cannot be solved in the same way.[8] For this reason, a rather cagey and odd kind of ontological attitude is usually taken with the collection of all the sets.

1.2. **Our Proposal.** But does this mean that the problem for sets is intractable? Just as with the case of the natural numbers, isn't it similarly obvious that we can talk about the collection of all sets? I appear to be doing it right now. I think the right way to understand what was happening in the previous section is that there is really only one kind of problem at play in the two examples. Moreover, in moving into the world of sets we haven't so much solved the problem as pushed it to the edges. So while we *"solve"* it in the locale of arithmetic, essentially the same problem pops up again in the world of sets. It may seem that the analogy breaks down but this is just because we tried to apply the exactly the same solution to the problem of sets, rather than one more appropriately suggested by the analogy.

So rather than attempt to solve the problem for sets using a set, we shall collect up the sets into a thing which is not a set, but a different *kind* of collection. There is a precedent for this in traditional set theory. Such an object is called a *proper class* [Kunen, 2006]. It is called a *class*, so we can distinguish it from a set; it is called *proper* because it is only a class and not a set as well. The move is similar to the approach of second order set theory [Shapiro, 1991]. Such theories, while philosophically somewhat fraught, are well-understood by mathematicians.

But there is no need to stop at the level of mere classes. What is stopping us from considering the collection of all of the classes? We seem to be doing that right now. Let a *family* be an arbitrary collection of classes. Then it is easy to arrange things so that there is a family of all the classes. But then we might want to know about all of the collections of families. This can go on ... indefinitely. For those familiar with set theory, this process should be familiar. At each stage of this game, I'm adding what is essentially the power-collection of the previous level. We are, more or less, repeating the process of construction of the cumulative hierarchy of sets.

So why not add indefinitely levels of this new hierarchy? Such a question has already been posed by Boolos who suggests that once classes are countenanced, then there is no reason not to continue [Boolos, 1998b].[9] Let us call this $SETS + SETS$ since we are, crudely speaking, *stacking* one notion of collection on top of another. To make formal sense of this, we might expand our language with a new collection relation η which extends the ordinary notion of membership: if $x \in y$, then $x \eta y$. The idea is that the $\in$ relation is sufficient to deal with all of the collections given by the ordinary cumulative hierarchy - the first $SETS$; while the η relation is intended to deal with collections from both the ordinary cumulative hierarchy and also the second hierarchy that follows it - so, $SETS + SETS$. Let us call objects from the ordinary cumulative hierarchy ($SETS$) $\in$-sets and those from $SETS + SETS$, η-sets. If an η-set is not an $\in$-set, then we shall say it is a *proper* η*-set*. To ensure that big collections are captured and thus the η relation properly extends $\in$, we demand that there is proper η-set which contains (via η) all of the $\in$-sets; and we call such an object a *universal* $\in$*-set*. Beyond this, we may demand that η behaves like $\in$ and thus constrain it with the usual axioms for ZFC with η replacing $\in$. The full details

8 Alternatively, we might take this problem as evidence that our collecting up of the natural numbers into a set is also problematic. Since the collecting failed for R, perhaps we should be suspicious of its abilities to collect the natural numbers too. This could then lead to an ultra-finitist position, which it outside the scope of considerations in this paper [Nelson, 1986].

9 Of course, his point is to head off talk of classes too. With regard to superset theories, he remarks, "I can't believe that any such view on the nature of "$\in$" can possibly be correct. Are the reasons for which one believes in classes really strong enough to make one believe in the possibility of such a hierarchy?" Our answer is yes, but given Boolos's nominalist tendencies and his paranoia about even set-sized entities, his reaction is understandable [Boolos, 1998a].

are not particularly important in this philosophical discussion, however, a detailed axiomatisation is provided in the appendix. We note that there is a proper η-set, R which contains all those $\in$-sets which are not $\in$-members of themselves. Thus, a version of the Russell set can be represented.

It should be noted that there is an easier way of formalising this, which is perhaps more traditional. Rather than using multiple membership relations, we might simply add a new constant symbol U to denote the ordinary cumulative hierarchy. Then U would contain all of what we have called the $\in$-sets and what we called the η-sets would be those x such that $x \notin U$. From a technical point of view, there is little to chose between these approaches beyond aesthetic considerations. Moreover, the latter approach is probably a little easier to glean at first blush. I've opted for the former approach as a matter of philosophical emphasis. We are not only trying to make the point that we can use set theory's $\in$ relation to describe the situation in which there are multiple universes - each congenial for mathematical practice. We are also trying to emphasise that this passage from universe to universe via indefinite extensibility involves a kind of ambiguity or indeterminateness in the very concept of membership. This will be discussed in more detail in Section 3.4.

But why should we be satisfied by $SETS + SETS$? We may want to talk about the collection of all η-sets or the collection of all those η-sets that are not η-members of themselves. This would lead us into what we might denote as $SETS + SETS + SETS$. We simply generate a further cumulative hierarchy over $SETS + SETS$. But why not go further still? Why not complete the process of adding consecutive new versions of $SETS$ indefinitely. We might call this $SETS \times SETS$ since we are, so to speak, adding $SETS$ many new versions of $SETS$ on top of each other. To express this formally, we might generalise the approach above by placing an index on the membership relation. We then write $x \in_z y$ to mean that x is a member of the collection y according to the membership relation index by z. These indices will be well-ordered so that if z precedes z', then $\in_z$ describes a universe which can be regarded as an initial segment of the one governed by $\in_{z'}$. Precise details of how we might axiomatise this are described in the appendix.[10] Again, we could also formalise this from the universe perspective by adding a function symbol, u, which takes an ordinal and returns a universe - in such a way that each universe is an extension on top of the previous one.

But why should we be satisfied by $SETS \times SETS$? There's no reason for us not go further again; for example, we might define something of the form $SETS^{SETS}$. Indeed, we can keep on doing this kind of thing indefinitely. This is the driving idea behind my proposal. The theories we've just proposed are merely examples of this idea. For the rest of the paper, we shall call these kinds of theories, *superset theories*.

1.3. **Guiding principles.** Before we move on to situate our suggestion and discuss some objections, let's take a step back and try to draw out the motivating principles that underlie an acceptance of superset theories.

First of all, there is a kind of closure principle in play.

> **Principle 1:** If a collection can be described (in an *acceptable* fashion), then it can be the value of a bound variable.

This principle gave rise to the problem: we deemed the collection of all sets to have been described in an acceptable fashion and thus, capable of being in the range of our quantifiers. The principle has historical precedent in Cantor's domain principle and Frege's problematic basic law V. I would emphasise, however, that Principle 1 is not stated with mathematical precision. This is intentional. I do not want to incorporate this principle into an axiomatic theory; indeed, I do not believe that it can be faithfully rendered in that manner. We

[10] Similar work has been done with regard to the truth predicate by Halbach [1995].

should also ask what is meant by acceptability here. We shall discuss this in more detail soon, but minimally it is some promise of consistency and more ambitiously coherence in a more philosophically satisfying sense.

In the previous section, we sketched a kind of interpretation for these superset theories, but this left open the axiomatisation of these new membership relations. I contend that exactly the same principles should be used to govern supersets as ordinary sets, so we should use *ZFC* (suitably adapted).

> **Principle 2:** If a collection-theoretic principle is good enough for ordinary sets, then it is good enough for supersets too.

Thus for example, the axiom of pairing should be true with regard to any membership concept.[11] This is a separate principle. Moreover, we should note that many philosophers become suddenly parsimonious when objects which are "larger" than sets are considered, so this is contentious. As such most of the forays into the world of supersets have stopped at the level of proper classes or even before that point. In response to this, I contend that it is hardly unreasonable to think that if an axiom like replacement is good enough for ordinary sets, then it should be acceptable for supersets too. Barring problems of consistency or coherence, I take it that the principles of *ZFC* are good for any notion of collection whether they be sets or supersets.

2. Traditional Responses to the Motivating Problem

Let's now consider a few traditional responses to our indefinite extensibility inspired problem. I won't claim the list is exhaustive, but it should give some idea of the lie of the land. We shall see that each response has non-trivial costs and that there are no clear winners at this time.

2.1. **Reject it as a problem.** This is probably the mainstream approach. Rather than indulging the concerns of the previous sections, we take classical *ZFC* at its word: it is *the* theory of collections. As such, if set theory does not say there is a set of certain kind, then there is no such collection. Thus, while it may appear that I was talking about the collection of all sets or the collection of all ordinals, this was actually some kind of illusion. We cannot let the collection of all sets be the value of a bound variable: there is no such collection.

While this may appear counter-intuitive, this is the bullet we must bite if we take set theory seriously as *the* theory of collections. Given the paradoxical nature of the subject matter, we might take it that our intuitions are somewhat impaired; and as such, we should treat this - not as an oddity - but rather as the core insight delivered by Russell's paradox. On the other hand, those who do not accept this view should be tempted to see this solution as blocking the problem by mere stipulation.

The response does have adherents. For example, the canonical set theory text books of Jech and Kunen both appear to uphold it [Jech, 2003, Kunen, 2006]. These textbooks do, however, admit what are sometimes known as *virtual classes*. For example, we are allowed to use a name V to refer to all of the sets. However, this is to be understood as mere convenience, a piece of metalinguistic trickery which can always be translated away. Most importantly, they do not permit classes to be the values of bound variables.

But we should be a little reserved here. Just because the standard mode of presenting set theory uses locutions which entail no metaphysical commitment to supersets and classes does not imply that users of that mode are not committed to the existence of such entities. One could believe in all manner of supersets and still elect to use

[11] The difficulty here is not so much the internal axioms of each membership relation but how different membership relations are related.

the modest mode of presentation so as to avoid controversy.[12] As we shall see in the second half of the paper, one can quite happily do anything of mathematical interest without the need to invoke classes or supersets.

2.2. **They're not sets, they're pluralities.** The next response we consider addresses the problem by arguing that the relationship between quantification and ontological commitment is not as straightforward as one might have expected. We first observe that natural language appears to admit more than one form of quantification. In particular, beyond the ability to say, *"there is a* dog," which can be comfortably represented in first order logic, we also seem to be able to say that *"there are some* ponies," which does not fit so comfortably into our ordinary logical apparatus. The point is usually pressed using the following famous example:

(∗) Some critics admire only one another.

This sentence cannot be straightforwardly modeled using ordinary first order logic. One way to deal with the problem is to bring in some set theoretic apparatus so that the initial quantification does not say that there is some critic, but rather that there is some set of critics. Boolos, however, regards this response as unnecessarily prejudiced against a perfectly good form of natural quantification [Boolos, 1984]. Thus, we should see this problematic sentence as evidence of an incompleteness in our theory of first order quantification. In response, Boolos introduces a new form of quantification which allows us to quantify over some critics without the need of set theoretic machinery. This is known as *plural quantification*. The purported advantage is that by using the logical apparatus of quantification to talk about a plurality of critics, we are not ontologically committed to the existence of a collection of critics only to the existence of the critics from that plurality. This claim is itself not without its critics, although this is beyond the scope of this paper [Hazen, 1993, Resnik, 1988].

Returning to the motivating problem, let's say we wish to talk about those sets which are identical to themselves; i.e., all of them. The adherent of plural quantification has a new move up their sleeve. When we appear to allow the collection of all sets to be value of a bound variable, we have two choices: it could be an individual variable, or a plural variable. Taking the first option they would end up with the trouble we saw in the first section, so it must be a plural variable: we are merely saying that there are some sets. Moreover, if we take up the claim at the end of the previous paragraph, then we are not committed to the existence of some collection of all sets, merely the sets in the plurality caught by the plural quantifier. In this way, we are supposed to avoid our problem since there is no ontological commitment associated with being the value of a plural variable.

This sits a little strangely with how I have articulated Principle 1 above. I attempted to remain somewhat quiet about ontological commitment and merely address the issue of something's being in the range of quantification. As I have described the matter, the adherent of plural quantification is avoiding ontological commitment while they still accept those sets which are self-identical can be the value of a bound variable.[13]

2.3. **Generality relativism.** Another response takes up a doubt that our quantifiers do as they purport to. Ordinarily when we say everything that is a dog is a mammal, there don't appear to be any problems. However, the generality relativist argues that phenomena like Russell's set and the problems of indefinite extensibility demonstrate that the range of our quantifiers is vexed. We should understand that our quantifiers do not really range over *absolutely everything* merely some determinate sub-collection thereof. Indeed, according to this

[12]For a more up front adherence to the rejection response, Mayberry's [1977] provides. Mayberry discusses a similar problem to the one we have been investigating. He observes that if we indulge in the effort to provide a metatheory for set theory, we appear to become stuck in a kind of regress. Mayberry's solution is to stipulate that set theory should be the theory upon which our spade turns.

[13]The adherent of this approach should see no problem in this. While a plurality of objects may be quantified over using a single variable, this does not of itself entail that the plurality is an entity over and above its members.

view, it doesn't really make sense to talk about *absolutely everything*. In actual fact, the range of our quantifiers can shift according to context. Views in this area are known as *generality relativist*. For a good discussion and defence of these views, see [Glanzberg, 2006] and [Uzquiano, forthcoming].

Returning to our problem of talking about the collection of all sets, the generality relativist tells us that the problem - if there really is one - is one of perspective. We take set theory seriously as *the* theory of collections, but we reject the idea that our quantification is absolute. Now when we set the problem up, we are trying to talk about the collection of all sets, in the sense of allowing that collection to be in the range of our quantifiers. Arguments like that of Russell's paradox, show us that this is not possible. However, this is only impossible according to the *initial* range of our quantifiers. It seems clear that there is some kind of collection corresponding to all of the sets captured by our initial quantification, so we allow ourselves to capture such a collection in a *new* range of quantification. From this new perspective, we are able to capture the collection of all sets according to the initial range of quantification. But of course, this doesn't give us the ability to talk about the collection of all sets according to the new range of quantification. If we want to do this, we must shift the range of quantification again using the same principle. So there is a sense in which the problem is merely deferred rather than solved.

2.4. **Logical revision.** A more radical response is to question, not merely, the metaphysics of quantification but the logical apparatus we use to illustrate our problem. The underlying idea of this approach is the thought that our problem highlights something defective in our reasoning rather than in our notions and theories of membership. A prominent and bold example of this approach is found in Weber's [Weber, 2010b,a]. Weber's goal is to develop a set theory which upholds naïve comprehension and extensionality and still provides a viable foundation for mathematics. His approach is both paraconsistent and dialetheic. When we encounter Russell's set we take the proof to inconsistency at face value: Russell's set is both a member of itself and not a member of itself. This is to be regarded as an insight rather than a hurdle. To keep things working smoothly, a paraconsistent logic is employed that blocks the argument that takes us from here to the conclusion that any sentence whatever is a theorem: triviality.

Returning to the problem of talking about the collection of all sets, Weber's use of naïve comprehension allows us to talk about this collection. Indeed it's a set! So unlike the previous approaches, we are allowed to both: admit this collection into the range of our quantifiers and maintain a unified approach to quantification.

The cost is, however, very high. Even if we countenance such a deep revision of our reconstruction of mathematical reasoning, it is still not yet clear whether such a move will actually provide an adequate foundation for mathematics. However, given that this approach is comparatively new on the scene and the difficulties are great, I think some reticence toward over-criticality is warranted. If a heuristically pleasing logic could be developed which provided a foundation for mathematics and dealt with the paradoxes, then I think we should be interested in such a project.

2.5. **Modal approaches.** This approach also involves a change in logical resources. However, rather than revising our logic, we expand it. The underlying idea exploits a relationship between infinity and possibility by, in some sense, re-envisaging Aristotle's notion of potential infinity [Parsons, 1983]. In these systems, we are permitted to talk about infinite collections, but absolutely unbounded collections whose ranks exhaust the ordinals remain merely potential. Indeed, the ordinals themselves are treated this way. For expositions of approaches along these lines, see [Linnebo, 2010, Studd, 2012, Hellman, 1989, Parsons, 1983].[14] A related

[14] Reinhardt's imagination based approach could also be thought of as modal [Reinhardt, 1974].

view is held by Andreas Blass how takes it that whenever we appear to refer to proper classes, we are *actually* referring to some V_κ [Blass, 2013, 2011].[15] There are many more collections that we *could* have quantified over, but we cut things off too soon.

There is a natural conceptual overlap between possibility and infinity that makes this approach philosophically very attractive. However, there are also costs. The most pressing issue is getting a metaphysical grip on the kind of modality used in axiomatising these systems.[16] Unlike the simple $S5$ interpretation usually given to metaphysical modality, modal set theories often require more complex systems somewhere between $S4$ and $S5$ in order to get a plausible modality for the *construction* of sets.[17] Of course, model theories can be provided for such logics which can be philosophically illuminating, but these can only be regarded as toy models for the modal set theories. This then leaves open the problem of how to bootstrap the insights gained from the toys up into the real worlds of modal set theory. That said, this is a problem shared, in some fashion or other, by many solutions on the market. Moreover, after further research it may turn out that such problems are least damaging in a set theory with modality.

2.6. **Property approaches.** This response to our problem takes seriously the idea that there is something which behaves a lot like the collection of all sets, but this thing is not a set: it's a property. Thus, we have the property of being a set and this property has as its extension all of the sets. Usually, properties are understood as the intensional cousins of the extensional sets. Thus, while any two sets with the same members are actually identical, we can have two properties containing exactly the same members. For a hackneyed example, consider the property of having a heart and the property of having a kidney. We are then supposed to see that these properties have exactly the same extension: the things which have hearts are exactly the same things as those with kidneys. However, they are not the same property. Thus, there is something more to being a property than being a set.

Turning to our main problem, we see that this gives us something like a collection of all the ordinals, but it's not a set. Indeed this articulation is a little infelicitous. There is a property of being an ordinal. We are thus to regard set theory as the tame domain of collections in which we can do mathematics. Versions of this approach to sets are defended by Hartry Field, Jc Beall and Tony Martin [Field, 2008, Beall, 2009, Martin, 2001].

There is something appealing here. It seems to give us a way around our problem and it deploys some relatively commonplace metaphysics to do it. However, the approach is relatively nascent and faces a couple of important challenges. First, we would like to know more about how properties work? An axiomatic theory would be the gold standard here, but even a sketch of informal principles would also be useful. Some work has been done here, although more work is required [Linnebo, 2012b]. More seriously, one might wonder why we would think that the property of being an ordinal had no extensional counterpart. In other words, why are the ordinals a *proper property*, so to speak. Surely there is a fixed collection of things which satisfy the definition of being an ordinal.

2.7. **Theories with a universal set.** Finally, we consider approaches which admit the existence of a set of all sets: a universal set. A particularly elegant example of this approach can be found in Forster's [2008].[18]

[15] κ in this case will be inaccessible and thus we have - as we'll define later - a natural model of ZFC.

[16] For some examples of attempts to get to grips with these problems, see Fine's [Fine, 2005] or Linnebo's [Linnebo, 2012a].

[17] It should be noted that Hellman's approach from [Hellman, 1989] avoids talk of *construction* and is able to work within the simpler modality of $S5$.

[18] It should be noted that unlike a great deal of Forster's work, this is not directly related to Quine's NF.

Forster employs a double barreled construction which can be thought of as generalising the ordinary approach of generating the sets via the cumulative hierarchy. In the ordinary construction we start with the empty set, which is V_0, and then take all of the subsets of that set to form V_1, which is just $\{\emptyset\}$. We then iterate this process through the ordinals. At each stage, Forster describes us as *lassoing* all the subsets of the previous level. Forster generalises the construction by allowing us to lasso not just subsets of a given level, but also their complements. Again we start with the empty set and we are able to lasso up the only subset of $\emptyset$, which is just $\emptyset$ itself. However, we are also able to lasso up the complement of this set; which in the case of the empty set, must be the universal set. We then iterate this process. The result is not the standard cumulative hierarchy - for starters it is not well-founded. However, it does contain a set which contains every set.

Returning to our focus problem, when we consider talk about the collection of all sets, Forster has a simple answer at the ready: the universal set. This approach involves no revision or augmentation of logic nor worries about the range of our quantifiers. Moreover, it is mutually interpretable with our ordinary set theory, *ZFC*, so nothing mathematical has been sacrificed.

However, we should also note a reservation. If we consider Russell's class, then we find that there is no set in Forster's system corresponding to it: there is no set of all those sets which are not members of themselves. This has a nice consonance with standard set theoretic approaches, which is arguably a good thing given the problems with this class are well-known. But similarly, there is no set of all well-founded sets, nor any set of all ordinals. These collections, on the other hand, seem completely reasonable and worth talking about.

Remarks. We can see from the above discussion that there is a wide market of candidate solutions to our focus problem. We are invited to revise our logic and our metaphysics in an effort to close off the theory of membership. However, each of the approaches above comes with an associated cost. Thus far, this field has no clear winners. In contrast, the solution offered in this paper is to extend set theory rather than revise it. We claim that there's nothing wrong with ordinary set theory, there's just more to the story. And while you can always tell more of that story, you cannot finish it. Thus, rather than offering a straightforward but disquieting solution to our problem, we are instead offering a way of learning to live with it.

3. Responses to Problems & Objections

We now discuss some problems and objections to the supersets proposal we have made in Section 1.2. We start with some quicker problems and then move on to more difficult challenges.

3.1. **It's not really a solution to our indefinite extensibility problem.** Our first observation is that we have not really solved our problem from indefinite extensibility so much as illustrated that the problem may be reinstated at a new level. Moreover, on the basis of this phenomenon's indefinite occurrence, we have taken this reinstatement as a kind of philosophical principle. Perhaps this is a problem for the proposal. We got here by complaining that Cantorian set theory did not really give us a solution to the problem of indefinite extensibility, it merely deferred it [Priest, 2002]; and now we are proposing an alternative which continues to suffer the same problem. This will be a theme throughout the remainder of this paper, but for the moment I remark that I think this is just what happens if you take indefinite extensibility seriously. I want to think of this approach as being modest in the sense of being *intellectually honest* about the pervasiveness of the problem. We are not so much trying to remove the problem as track it more faithfully.

But I think we can also say something deeper and more interesting at this point. ~~At heart,~~ Principle 1 tells us that if a collection can be described, then it can be the value of a bound variable. At heart, this is another manifestation of indefinite extensibility. Given a *definite totality* in the sense provided by a particular membership relation, we can *by reference to that totality* form a *larger totality* via a new membership relation.[19] The supersets outlook takes a deeper interpretation of indefinite extensibility and turns it into a feature of the underlying philosophy.

3.2. **Isn't this just generality relativism?** We might also worry that our suggestion is just another form of generality relativism. However, there is an important difference. While the generality relativist claims that the range of our quantifiers is subject to shifts according to context, our suggestion is that the membership relation itself is subject to change: we were thus mistaken in talking about *the* membership relation. This idea has been explored in Williamson's [1998]; we might call it *membership relativism*.

There is also, however, a close affinity with generality relativism. Consider a system of domains of quantification as used by a generality relativist. From each of these domains of quantification, we may define a particular notion of membership whose field is that domain. In this way, we can extract a plurality of membership relations in accord with the proposal of this paper. Similarly, we may move in the other direction by taking a plurality of membership relations and extracting a particular domain of quantification from each of them. Thus, we might say that the two approaches are logically or mathematically equivalent. However, to the extent that the range of our quantifiers is of metaphysical significance, the philosophical differences are of great importance. The best way to illustrate this is to consider the debate about absolute generality and the question of whether our quantifiers range over absolutely everything. If we say yes, then we face the problems of indefinite extensibility; if we say no, then we face Williamson's paradox from [2003]. But there is a third way: if we take up membership relativism we are afforded the opportunity to say nothing at all.

To see this, we claim that membership relativism is not only compatible with the generality relativist's shifting domains of quantification, but also compatible with the fixed domain of quantification adopted by the generality absolutist. This might seem strange given that we've just noted above that we may associate a particular domain of quantification to each of the membership relations in a superset theory. However, there is no compulsion for the superset theorist to admit multiple domains of quantification: they can get by with just one. Thus for example, consider the move between ordinary set theory with just $\in$ as its membership relation and a theory of $SET + SETS$ with say $\in$ and η. All that changes is that we add some new axioms to deal with η.[20] While it's true that more things are related by η than by $\in$, there is no reason to think those objects weren't out already out there when we only used ordinary set theory: we just didn't have much to say about them. Thus, there is no reason to think that the range of our quantifiers shifted in this move. The superset theorist can accept a single domain of quantification.

[19]By taking limits of sequences of membership relations, as in the $SETS \times SETS$ example, we can make theories which appear to capture this phenomenon, but which are still prey to Principle 1. No matter what, something seems to be left out. It is for this reason that we call Principle 1, philosophical rather than mathematical.

[20]See the Appendix for a detailed axiomatisation of $SET + SETS$. It is, however, worth observing a certain deviance in the axiomatisation. Consider the proper η-set, $V^{\in}$, of all the $\in$-sets and the proper η-set, $On^{\in}$, of all the $\in$-ordinals. $V^{\in}$ and $On^{\in}$ are objects in our theory and as such we might expect that the axiom of pairing for $\in$ to give us some z such that $V^{\in} \in z$ and $On^{\in} \in z$. This cannot occur in the theory of $SETS + SETS$; thus, the full axiom of pairing fails. While there will be some y such that $V^{\in} \eta y$ and $On^{\in} \eta y$, some damage to our intuitive understanding of pairing has been done.

3.3. **Hasn't this been done before?** One might also get the feeling that the view proposed here is somewhat familiar and unoriginal. I think there is something right about this inclination, but there is also an aspect in which it misses something very important. To draw this out, let us look at two plausible precedents for the view proposed above.

3.3.1. *Zermelo's "meta-set theory"*. Our proposal certainly shares something with Zermelo's thoughts in "*On boundary numbers and domains of sets"* [Zermelo, 1976]. Very briefly, Zermelo outlines a view of set theory which is formulated in a second order logic and which speaks of multiple normal domains or models in which each of the usual axioms hold. We see that

> *"what appears to be an "ultra non- or super-set" in one model, is, in the succeeding model, a perfectly good, valid set with both a cardinal number and an ordinal type, and is itself a foundation stone for the construction of a new domain."*

This seems to be in accord with Principles 1. and 2. Zermelo is saying that we can always move to a new model in which classes become sets in the new domain - this is like our move of taking up new versions of the membership relation, each extending the previous membership relation. Moreover, the ordinary axioms are maintained as we move from one model to the next - as required by Principle 2.

But there is, I think, a difference too. Zermelo also argues that his framework has provided a solution to our problem from indefinite extensibility - or as he calls it, the ultrafinite antinomy. We are not making that claim, at least not in the sense of having provided an axiomatic system. While I am reluctant to claim to know how to correctly interpret Zermelo, he appears to argue that he has provided (at least a step toward) a formal foundation of this.

> *"The existence of an unbounded sequence of boundary numbers must be postulated as a new axiom of "meta-set theory" ..."*

From our perspective, this seems misguided. We claim that any meta-set theory is subject to the same problem: a kind of indefinite extensibility revenge. It is here that we come apart.

3.3.2. *Grothendieck universes.* More recently, mathematicians have taken to using Grothendieck's toolkit, in particular Grothendieck universes, in their work: a famous example being Wiles' proof of Fermat's last theorem [McLarty, 2010]. At heart, a Grothendieck universe is a collection of sets closed under properties useful in the construction of mathematical objects in, say, algebraic geometry.[21] When working with these universes it is often convenient to assume that every set is an element of some universe. This is known as the *universe axiom*.[22]

Once again, we see something in common with our proposal. Principle 1. appears to be upheld since for any set we care to take, there is a universe containing it. Given that universes are themselves sets, it is easy to see that a chain of universes each extending the previous one is generated, much like the picture we gave above. Moreover, Principle 2. is satisfied since each universe is closed under the useful properties. So it is clear that

[21] A full axiomatisation can be found in [Bourbaki, 1972].

[22] We should note that the universe axiom is equivalent to the existence of unboundedly many strongly inaccessible cardinals, so the theory here is noticeably stronger than *ZFC*. However, generalising Cohen's trick with standard models, Feferman has shown that theory with seemingly cosmetic differences to Grothendieck's is equiconsistent with *ZFC* [Cohen, 1966, Feferman and Kreisel, 1969]. As we shall see, this is not surprising given that inaccessible cardinals are comparatively weak and are not known to have any tangible effects on combinatorial questions of palpable mathematics. In particular, McLarty has shown that Fermat's theorem can actually be proven in finite order arithmetic, which is very much weaker than *ZFC*.

we have an example of a superset theory; indeed something very close to what we have called $SETS \times SETS$ above.[23]

However, there is also an important philosophical difference between the underlying projects, which is most easily highlighted if we consider Principle 1. and the universe axiom. While the universe axiom tells us that every set is contained in some universe, it does not tell us that all of the sets are contained in any one universe, or that all of the universes are contained in a universe. These kinds of question outstrip the Grothendieck universe framework. I do not want to suggest that this is evidence of a fault in the framework - at least no more a fault that in any other framework - this kind of question is simply outside its scope. Nonetheless, the universe axiom fails to properly satisfy Principle 1: something is still left out. Of course, I haven't provided a formal theory that does this either; and indeed, I am suggesting this would not be possible. So while the theory of Grothendieck universes provide a plausible superset theory candidate, it could not be construed as solving the underlying indefinite extensibility problem. It's just more grist for that mill.

The real point here is one of emphasis and motivation. Grothendieck was motivated to provide a congenial toolkit for working mathematicians, where we are motivated by a generalisation of the problem of indefinite extensibility. There is no problem using a theory which allows us to talk about a system of universes, each closed under nice properties: we do this in ordinary set theory all the time. There is, however, a difference in explicitly suggesting that the multiplicity of these universes is witness to something fundamentally incomplete about the notion of membership itself. My thought is not so much to highlight the usefulness of supersets in mathematics - I don't think they are, but rather to use them as a vehicle to draw a difficult philosophical problem out into the open.

3.4. **Can there really be more than one notion of set?** In admitting a plurality of collection relations rather than just one, there is a reasonable niggle with regard to the intelligibility of the supersets proposal. Surely there is only one notion of membership to be had. How can we even make sense of a plurality of them? Indeed, what are supersets, if not just more sets, but now under some other name?[24] To answer this question, we appeal to an analogy with formal theories of truth. In essence, I would like to argue that a respectable approach to truth theory finds itself in much the same position as the set theorist: they find themselves with a plurality of truth predicates where their default intuitions suggest there should be just one. While such a plurality of truth predicates is bound to be disconcerting, given that the other options involve deep revision of logical resources, we can be at least partially satisfied in throwing a predicate prone to paradox to the wolves rather than abandoning reason itself. We then claim that there is a plurality of membership concepts in just the same way that there is a plurality of truth concepts; it's not what we would have hoped for, but we can live with it. This response could well be disappointing to the reader since we are not providing a positive conception of this pluralistic picture, merely claiming that other respectable positions are in the same boat. This, however, is entirely in line with the spirit of this paper in that we aren't so much offering a bold new solution as figuring out the right way to bite a difficult bullet.

Let us first consider a relatively simple example of a formal theory of truth. In [Tarski, 1956], Tarski demonstrated how one might provide an axiomatic characterisation of the truth predicate for a theory like arithmetic.

[23]Reinhardt and Ackermann also developed set theories along similar lines, see [Reinhardt, 1974]. While there are important technical differences between these approaches, they share the same philosophical shortcomings from our superset theory point of view.

[24]I thank the referee for this particularly succinct way of putting the problem.

Very briefly, we expand the language of arithmetic $\mathscr{L}_{Ar}$ with a predicate "T_0" whose intended intended interpretation will be the set of codes of true sentences of arithmetic. Call this expansion $\mathscr{L}_{Ar}(T_0)$. The resultant theory, $T(PA)$ is an extension of PA and is in common use in mathematical logic.[25] However, such truth definitions are restricted in an important way: they are only informative about the truths of sentences from the language $\mathscr{L}_{Ar}$, not the rest of the sentences from $\mathscr{L}_{Ar}(T_0)$. Thus, for example in $T(PA)$ we can show that $0=0$ is true.[26] However, we cannot show in $T(PA)$ that $T_0\ulcorner 0=0\urcorner$ is true.[27] The reason for this is simply that the clauses characterising truth in $T(PA)$ only concern sentences of $\mathscr{L}_{Ar}$.

But perhaps this is just an artefact of the theory $T(PA)$; perhaps we can fix this glitch and provide a theory that allows can establish the truths of sentences which make use of the truth predicate and not just those from $\mathscr{L}_{Ar}$. This leads us straight into the liar paradox. Suppose we provide a theory Δ extending $T(PA)$ in the language $\mathscr{L}_{Ar}(T_0)$ which ensures that T_0 is a reasonable truth predicate for the full language $\mathscr{L}_{Ar}(T_0)$. Then for every sentence φ from $\mathscr{L}_{Ar}(T_0)$ we would expect that:

$$\Delta \vdash T_0\ulcorner\varphi\urcorner \leftrightarrow \varphi.$$

Then since Δ extends PA, we can use the diagonal lemma to obtain a sentence λ such that:

$$\Delta \vdash \lambda \leftrightarrow \neg T_0\ulcorner\lambda\urcorner$$

which leads directly to a contradiction.

The argument here is structurally analogous to the argument we used to demonstrate that the existence of R (the set of all sets which are not members of themselves) led to contradiction [Priest, 2002, 1994, Lawvere, 1969]. Moreover, the underlying problems and upshots are also tightly related. In the case of truth, we aimed for a *single* truth predicate that could be meaningfully applied to any sentence whatsoever. The argument above can be construed as showing that no such truth predicate is available. In the case of set theory, we aimed for a *single* notion of membership that could be applied in the context of any collection whatsoever. We then take it that our argument about R showed that no such membership relation is available.

But what then if we want to say that $T_0\ulcorner 0=0\urcorner$ *is* true? Tarski provided a standard approach to this problem which has been thoroughly developed in [Halbach, 1995, 2011]. Rather than using just one truth predicate, we expand our language further with another predicate "T_1" and call the resulting language $\mathscr{L}_{Ar}(T_0,T_1)$. The intended interpretation of T_0 remains the same, but the intended interpretation of T_1 is the set of codes of true sentences from $\mathscr{L}_{Ar}(T_0)$ and not just $\mathscr{L}_{Ar}$. The resultant theory is not in common usage, but it is straightforward. Moreover, it will be able to prove that $T_0\ulcorner 0=0\urcorner$ is true and indeed it can easily be seen that the interpretation of T_1 will extend that of T_0.

In essence, this is the superset approach applied to truth. We are unable to get all that we wanted from T_0, so we add a new truth predicate T_1 to go further: this is much the same move as we took to get from $SETS$ to $SETS+SETS$. Of course, we bump into much the same problems. In $SETS+SETS$ we weren't able to talk about the collection of all η-sets. Similarly we are not able to show that it's *true* that $T_1\ulcorner T_0\ulcorner 0=0\urcorner\urcorner$. Of course, this can then be addressed by adding another truth predicate, say T_2. Indeed we can add infinite sequences of

[25] An explicit axiomatisation can be found in [Halbach, 2011].

[26] More formally, $T(PA)\vdash T_0\ulcorner 0=0\urcorner$ where $\ulcorner\cdot\urcorner$ is a coding function.

[27] More formally, $T(PA)\nvdash T_0\ulcorner T_0\ulcorner 0=0\urcorner\urcorner$. Indeed, depending on the coding function used, we may be able to show that $T(PA)\vdash \neg T_0\ulcorner T_0\ulcorner 0=0\urcorner\urcorner$.

new truth predicates just as we did in the case of $SETS \times SETS$. There is nothing stopping us going further and indeed there'll always be another troubling sentence that suggests that we should.[28]

So we've now seen an approach to truth which is motivated by similar principles and issues as superset theory. However, given some more recent research efforts into truth theory, one might wonder if such hierarchical approaches have been superseded. Leading candidates include methodologies taking up non-classical logics and approaches which are developments upon Kripke's fixed point techniques.[29] I suggest that we set these non-classical approaches aside, not because of any fatal deficiency, but rather because they form just one section of a market that is some distance from consensus. Given the analogy we are pursuing, we should see these programmes as close cousins to the non-classical set theorists we saw in Section 2.4.

On the other hand, truth theories developing our of Kripke's [Kripke, 1975] have a different flavour. Rather than attempting to revise our logic to obtain the so-called T-schema, Kripke provides an inductive construction which works its way up deciding the truth and falsity of sentences with ever deeper embedded truth contexts.[30] This inductive approach provides us with a means of determining the truth of sentences from a language like $\mathscr{L}_{Ar}(T_0)$ regardless of how often a truth predicate is applied to a sentence. In this sense, it could appear to provide a solution to our problem and thus sidestep our analogy. This thought is misleading. In fact, the analogy can be repeated here and it is arguable more faithful at this level. To see this we should first observe that the analogous achievement of a type-free theory of truth had already been achieved by set theory. While Kripke's theory of truth allows us to consider iterated truth contexts without needing to label the truth predicates; we can see that set theory allows us to reason with iterated collections without the need for a labeled theory of types [Church, 1940]. In other words,

> Kripke's theory of truth *is to* Tarski's hierarchies
>
> *as*
>
> Set theory *is to* the simple theory of types.

Thus, if we are looking for a strong analogy between truth theory and set theory, we are even better off if we focus on Kripke's theory. The second point we need to note that Kripke's construction also suffers a serious flaw: it doesn't decide every sentence to be either true or false. For example, the liar sentence fits into such a *gap*. Thus, since the liar sentence isn't true, we might expect that we could say that this very fact is true; i.e., that it is true that the liar sentence is not true. Kripke's theory does not permit this: this is often known as a revenge problem [Leitgeb, 2007]. However, there is a way around this. We can introduce a new truth predicate "T_1" into the language and run Kripke's construction again.[31]In the resultant construction, we are able to see that $\neg T_0 \ulcorner \lambda \urcorner$ is indeed true in the sense that its code is in the extension of T_1. This is essentially superset theory

[28]In [Halbach, 1995], Halbach explores theories involving transfinite sequences of truth predicates. He bounds them at ω_1^{CK} the supremum of the recursive well-orderings, however, the only reason to stop there is a bound on what a reasonable language should be like.

[29]Leading proponents of non-classical approaches include Priest [1979], Beall [2009] and Field [2008]. Rather than admitting an indefinitely extending hierarchy of truth predicates, these approaches revise our logical resources in order to avoid the difficulties of the liar paradox. The dialetheists among them, go so far as to argue that the liar sentence is both truth and false.

[30]The resultant fixed point construction can be understood through the strong Kleene logic, which is non-classical. So there is a sense in which the two approaches enjoy some overlap. However, the underlying methodology of Kripke's approach is rooted in the inductive construction, not the logic. Indeed there are fully classical variations which have been developed along the same lines [Leitgeb, 2005].

[31]This has been explored by Glanzberg in [Glanzberg, 2004], however, the construction is relatively straightforward. We use the inductive construction to define an extension for T_0 and then fixing that interpretation we run Kripke's construction again to get an extension for T_1. It should, however, be noted that there are alternatives to introducing a hierarchy of truth predicates. In response to revenge problems: Cook has thoroughly investigated the use of a transfinite hierarchy of truth values; and Schlenker has investigated transfinite hierarchies of negation operators and their relationship with transfinite truth approaches [Cook, 2007, Schlenker, 2010]. Nonetheless, the underlying theme here is that in order to respond to revenge a semantic concept ends up being stratified in a manner contrary to our initial expectations.

applied to the truth theory. And of course, just as we saw with the Tarskian hierarchies, a new liar sentence can be defined for T_1 which is in its gap. This can be addressed by adding yet another truth predicate, and so it goes on. Just as superset theory must always countenance another membership relation, the hierarchical approach to truth always allows for another truth predicate.

Hierarchical approaches to truth are exposed to an analogous problem to that faced by the supersets approach: what satisfies these new truth predicates if not just more true sentences? The answer in this case is that the standard classical picture of truth leads to a fragmented hierarchy of truth predicates perpetually approximating each other. There is no total theory of truth; and analogously with regard to the problem of this paper, we say that there is no total theory of collections.

3.5. **Deferral to mathematical expertise.** I would now like to consider a different line of objection to the view I have proposed. As opposed to the objections considered above, I believe that something in this area provides the best explanation as to why - in practice - the supersets proposal is so rarely considered. We shall argue that on a variety of construals this objection is not compelling. At its crudest, the objection is simply that:

> Mathematicians do not take this approach, so neither should philosophers.

At face value, there is something to this. Most mathematicians, in particular set theorists, assiduously avoid talk of proper classes, only admitting them as a *façon de parler* [Jech, 2003]. Moreover, if the experts about sets - i.e., set theorists - avoid talk about classes and supersets, then surely we should take this seriously.

We shall consider a few ways of refining the objection. First, we consider a kind of naturalistic argument. Mathematicians are the experts about mathematics and since set theory is a part of mathematics, we should *defer* to them. Second, we consider whether our talk of superset theories is *intelligible*. Given that as a matter of fact we have limited mathematical experience of supersets, we may have reason to doubt their coherence. Finally, we shall consider whether superset theories are *dispensable* and in such a way that we are warranted in ignoring them. We shall conclude that mathematicians have good reason to hold the views they do, but these reasons are less than philosophically compelling. The ensuing discussion is also of independent interest as it exposes a kind of rift between the respective goals of philosophical and mathematical research programmes with regard to set theory.

3.5.1. *Should we defer to mathematicians?* If we consider the landscape of the average mathematics department, set theory is, at best, a minor player. Most mathematicians have, at most, a cursory knowledge of set theory; while they could happily define an infinite sequence of functions for use in analysis, they would be hard pressed to describe the principle of transfinite recursion: it is simply not something they need to be able to do.

The value of set theory for mathematics is in the provision of a common underlying framework in which all everyday mathematics can be seen to take place. A set theory like *ZFC* is able to provide both an *ontology* of spaces within which mathematics can be done and a *proof theoretic* lever by which its theorems can be demonstrated. Everything in the mathematics department (with the exception of set theory itself) can be comfortably accommodated by it. It is in this restricted sense that *ZFC* successfully provides a foundation for mathematics. But this does not mean that the everyday mathematician needs to be familiar with it.

Most mathematicians would not notice if set theory was expanded, contracted or replaced by a completely new foundation. They would only care if such a change required them to re-think their ordinary practice. As such, there seems to be little point in deferring to a mathematician with no expertise in set theory on the subject of supersets: the question has no impact on them. The augmentation of ZFC by a theory of supersets will not take away any theorems and, as we shall see later, the theorems it could add are likely to be thought as having merely logical value.

But this still leaves the set theorists who generally have no truck with proper classes, let alone the supersets we are proposing. In this regard we make a further observation. Set theory is also good for mathematics because it pushes difficult philosophical questions to the very edges of its territory. While there may be problems for talking about the set of all sets, within the vast realm of the cumulative hierarchy, paradox is pleasingly absent. Set theory, as a mathematical discipline, can develop its rich tapestry of combinatorial problems without fear of treading into the lair of paradox.

But we may well ask if this is the whole story. While ordinary set theory provides a satisfying foundation for almost all of mathematics, we have seen that it is incapable of collecting up all the objects of its concern (i.e., sets) and manipulating them as a single collection. There is a gap in the foundation which set theory provides and set theory itself is located there. The cost of a paradox free foundation for ordinary mathematics is a lack of foundation for set theory itself. The superset theories we have discussed above promise to fill this gap albeit in an ongoing and perpetually deferred fashion.

So should we defer to the ordinary practice of set theorists on this matter? I think we should not, but with a couple of reservations. The set theorist's ordinary practice is, like any mathematician's, focused on the development and solving of problems. It is not, in general, focused on the problem of providing a philosophically satisfying foundation for mathematics. For the former purposes, ZFC is an ideal tool. But for our more lofty and perhaps impractical philosophical goals, this gap presents a more serious problem.

As to my reservations, I think we ought, first, demand that this talk of superset theories is coherent, at least for the purposes of mathematics. Second, we should require - just as we did for the ordinary mathematician - that superset theory has negligible impact on ordinary set theoretic practice. For example, we should be wary of a substantial expansion or reduction of the collection of theorems. We shall see deal with these issues in the following two sections.[32]

3.5.2. *Are supersets unintelligible?* Given that superset theories are unorthodox we may have cause to worry about their intelligibility. We see something like this in [Field, 2008] in the context of attempts to provide an absolute definition of truth. We first note that, in general, by augmenting our ontology with a layer of classes,

---

[32]We should also note that some set theorists are not so adverse to proper classes and indeed some important results are difficult to state without their aid. For example, Kunen's theorem tells us that there can be no non-trivial elementary embedding from the universe to itself. Without saying too much about its content, we may note that it has both philosophical and mathematical importance: it provided a devastating blow to Rheinhardt's large cardinal programme; and it provides a regularly used tool in the theory of large cardinals. However, an elementary embedding - albeit non-existent - from the universe to itself is so large that it can only be represented by a proper class. Thus, it seems that in order to even state the result, we need to move into the world of supersets. That said, there is a way of stating the essential content of this theorem without using classes. The important information about the embedding can be coded into a set and the following theorem suffices:

Theorem. *There is no* $j : V_{\lambda+2} \prec V_{\lambda+2}$ *for any* λ.

Nonetheless, there does seems to be something more natural about its proper class form. Should a set theorist prefer the class form without irony, I think we should accept that they are adopting a form of superset theory. Overall, this is good for an argument supporting superset theory: even if we adopt a naturalistic outlook there are set theorists to whom we may defer. However, there is also room to be somewhat reserved about this kind of evidence in that this says little about whether supersets were really required in these cases. Moreover, these forays into superset theory generally flounder at the level of proper classes or, at most, a level or so above that.

Other plausible examples of superset theory include: Easton's use of *class* forcing to code just any reasonable relationship we like between the $\beth$s and the $\aleph$s; and the Σ_2 well-ordering of mice - which exceeds the length of the ordinals - used in the construction of the core model K.

we can define the truth predicate for some base theory of the original ontology. We simply take the smallest class of sentences closed under the usual composition clauses [Wang, 1952]. So we can, for example, define a truth predicate for arithmetic by augmenting our ontology with classes of natural numbers and providing a weak comprehension principle. But what do we do when we come to the theory of sets itself?

If we allowed ourselves a layer of proper classes we can define truth for *ZFC* in, say, *NBG*.[33] But then we would lack a truth definition for the class theory. If we then wanted a truth theory for class theory, then we would need a further layer of families over and above the classes. This would then allow us to define truth using the same approach as before. The game then goes back and forth: add a layer of collections, define truth, add another higher layer of collections, define truth, ... Field uses this illustration to motivate the claim that truth *simpliciter* cannot be defined. The underlying insight might be summarised as follows:

> **Insight:** There is no *uniform* means of defining truth; there is no unique perspective from which truth can be defined.

This is in accord with our analogy between truth and superset theory, however, it does not rule out the possibility that for any particular theory, there is another theory in which a truth definition for it may be provided. Nonetheless, Field goes on to suggest that this use of **"**super cool entities**"** is somehow a little odd, embarrassing and even dangerous. Field states:

> *It is sometimes said that Tarski showed that a truth predicate for a language $\mathscr{L}$ is always definable, though only in a more powerful language $\mathscr{L}^*$. But this seems to me an extraordinarily contentious claim.*

It does not, however, seem particularly contentious from the point of view of the superset theories we have been discussing. Following our first principle, we can always get to a new layer of classes and then define a truth predicate using the technique described above. But perhaps supersets themselves are risky? The chief worry is that a superset theory may be inconsistent. We shall demonstrate that they have good reason to think that they are not. But before we do this we should note that there is something substantive in this claim. There are, in fact, theories which cannot be extended with classes to represent all the collections of objects of the original theory.

Example. Let $T = PA + \neg Con(PA)$; i.e., the result of adding to Peano arithmetic the statement that Peano arithmetic is not consistent. By Gödelian considerations, T has models. Let $\mathscr{M}$ be an example of such a model. Now let us expand our language by adding a layer of classes which will be accommodated by a new sort of variable. Augment T with the full second order comprehension axiom schema and amend the induction schema to take classes rather than formulae. Call the result T^+.[34] Then $T^+ \vdash Con(PA)$; thus T^+has no models and $\mathscr{M}$ cannot be expanded with a layer of classes.

The upshot of this example is that there is something to worry about and the fact that superset theory meets the challenge is further reason to take it seriously.

[33]This was first established by Wang in [1952]. The truth predicate is defined here in the quite weak sense that we can prove the T-schema for any sentence of the (truth-free) fragment of the language of set theory. Interestingly, *NBG* is a conservative extension of *ZFC*, so one might be concerned that the truth predicate would permit a consistency proof of *ZFC* to be conducted. However, *NBG* is not strong enough to carry out the further argument, where a theory like *MK* would suffice, although this is overkill [Wang, 1952]. For example, consistency can be established more economically by admitting class terms into the Replacement and Separation axioms. Similar remarks apply in the domain of arithmetic where ACA_0 can be used to define a truth predicate while consistency cannot be established.

[34]Shapiro [1991] shows how to do something very similar.

We now sketch some arguments for why superset theories are, in general, consistent and more. Recall that in order to demonstrate the consistency of a theory, by soundness, all we need to show is that the theory has a model. Indeed, we shall go further and show that not only do superset theories have models, but with very little cost we can produce models with very pleasing and natural properties.[35]

To do this, we shall make use of the notion of an inaccessible cardinal. This is a generalisation of the ordinary notion of infinity witnessed by the set of natural numbers ω. If we remove the axiom of infinity from ZFC, we can only prove the existence of finite sets. However, once we add it we are able to prove the existence of a wealth of larger and larger infinite objects. In the context of ZFC, the first inaccessible cardinal κ is an upper bound on the size of sets whose existence can be proven using the tools of set theory.[36] A little crudely, one might think of the axioms of set theory as a kind of machine that allows us to create larger and larger objects: sets. For example, the axiom of replacement tells us if we input a set of arguments into a definable function, the outputs resulting from this will also form a set. But given any particular axiomatisation, there will be some limit to the size of object which that machine can build. Thus, κ is so large that we can never define a function which reaches it. Similarly, in the case of the natural numbers, we can define all sorts of machines which can be used to construct enormous and complex numbers. But all of them will be bound by ω. So, loosely, we might say that

> ω is to the *theory* of natural numbers as κ is to the *theory* (ZFC) of sets.

Now since the first inaccessible κ is so large, if we consider all of the sets which are no *larger* than κ,[37] we get a *natural* model V_κ such that $V_\kappa \models ZFC$ [Kanamori, 2003, Zermelo, 1976]. Analogously, V_ω is a model in which ZFC without the axiom of infinity is satisfied.

Let us say that a (set-sized) structure $\mathscr{M} = \langle M, \in \rangle$ is a *natural* if:

- $\mathscr{M}$ is *supertransitive* (i.e., if $x \subseteq y \in M$, then $x \in M$ - i.e., M is closed under the *real* subset relation); and
- the set λ ordinals of $\mathscr{M}$ (i.e., $o(\mathscr{M})$) is *regular* (i.e., there is not function $f : \alpha \to \lambda$ where the range of f is unbounded in λ).

We say that such a model is *natural* since the most powerful axioms of the set theoretic machinery, powerset and replacement, are doing what they are supposed to do.[38] The powerset axiom tells us that given any set, the set of all of its subsets must exist. By demanding that $\mathscr{M}$ is supertransitive, we ensure that $\mathscr{M}'s$ powerset is the real powerset; i.e., it contains all subsets as seen from the outside. The replacement axiom is intended to describe the situation where every function with a set for a domain must have a set for a range. By insisting that the ordinals of $\mathscr{M}$ are regular, we ensure that every function whose domain is a set in $\mathscr{M}$ has a range which is a set in $\mathscr{M}$, regardless of whether f is in $\mathscr{M}$ or not.[39]

---

[35] Less natural models are offered in the Appendix which show that a couple of examples of superset theories are actually equiconsistent with ZFC.

[36] More strictly, ZFC does not allow us to prove the existence of any object whose transitive closure has cardinality greater than κ.

[37] Strictly, those x whose transitive closure has cardinality less than κ.

[38] We should note that a model's being natural is still quite a weak requirement in that we have said nothing about what sentences are true in such a model. For example, if κ is inaccessible then the set of subsets of κ of cardinality $< \kappa$ provides the domain of a natural structure, although it is not our real focus. For that we also need to know that our theory is satisfied there. Naturalness is being used here to ensure that some of the desirable properties of a model of set theory, but which are beyond the expressive capacities of first order logic, are still captured.

[39] Of course, another seemingly reasonable requirement of naturalness would be that the model of ZFC should exhaust the ordinals. That is, after all, what it is intended to do. However, if we make this move then we could not provide models of the superset theories since they

Now if we add another layer of sets to V_κ we get $V_{\kappa+1}$ which can be used as a model for a theory of proper classes like $ZFC2, MK$ or NBG [Shapiro, 1991]. Essentially, we let the elements of V_κ play the role of the sets of the theory; and we let the elements of $V_{\kappa+1} \backslash V_\kappa$ play the role of the proper classes. Although it does not satisfy the definition above, it is natural in the sense that the set theoretic machinery works as it is intended and since $V_{\kappa+1}$ contains all of the subsets of V_κ we have representatives for all of the classes. Going further, if we then consider the next inaccessible cardinal, then we get a natural model for $SETS+SETS$. If we consider the first inaccessible limit of inaccessible cardinals, then we get a natural model of $SETS \times SETS$. See the appendix for a more precise description of these results.[40]

It should be clear that each of the superset theories we have been considering can be shown to be consistent using a natural model. Thus at the minimal level required for mathematical coherence, we can show that these theories are coherent. This does not fully answer deeper complaints about the philosophical coherence of superset concepts. However in conjunction with our earlier remarks regarding the analogy with truth, one might claim that in using natural models, we are approaching deep coherence in that the models are as close to our expectations as we could hope for in this framework. The main point, however, is that worries regarding the mathematical coherence of these theories are unfounded.

3.5.3. *Are supersets dispensable?* Finally, we consider whether theories of supersets add anything to set theoretic research as a mathematical discipline. We shall argue that, from the mathematical perspective, superset theories are dispensable. We first make some general observations about the research landscape. Contemporary set theoretic research might be summed up as the study of problems involving infinite collections. One of the main tools for this kind of investigation are large cardinals. As we saw above, these are cardinalities which *ZFC* is not capable of demonstrating existence of. An inaccessible cardinal is one of the smallest members of this class. But not all large cardinals are mathematically *interesting*. An interesting large cardinal is one whose existence boosts the strength of the set theory in such a way that more, interesting problems can be solved. The kind of problems we have in mind are not *recherché*: by adding axioms asserting the existence of certain large cardinals, problems about real numbers can be solved.[41] Moreover, there is some consensus that this phenomenon provides us with evidence as to the significance - if not existence - of these large cardinals. A large proportion of contemporary set theoretical research is devoted to research in this area. We shall illustrate this below.

are intended to go beyond the ordinary ordinals provided by *ZFC*. This is just a limitation of *ZFC* to furnish the ontological resources for a thoroughly natural model of a superset theory. I think the right way to understand this situation is that this is as much naturalness as *ZFC* can accommodate.

[40]Indeed, this insight gives us an indication of how we might form a simpler axiomatisation of the superset theories. Rather than using a multiplicity of membership relations each extending their predecessors, we might just start with the axiom of infinity that would give us a natural model for these membership relations. All of this can be done within the language of $\mathscr{L} = \{\in\}$. The multiplicity of subsidiary membership relations can then be recovered afterward.

[41]The fact that the existence of such *prima facie* large objects increases our theoretical leverage on problems in the more worldly field of analysis is one of the more fascinating features of set theoretical research; a feature which currently lacks thoroughgoing philosophical explanation and understanding.

The smallest cardinal useful for such purposes is known as a measurable cardinal.[42] Like an inaccessible cardinal, it can also be understood in analogy to ω, but not as a limit of a process; rather it is the smallest cardinal, which has a certain useful and interesting property shared by ω. Moreover, there is a sense in which a measurable cardinal cannot be reached from below by merely *simply* increasing the power of the machinery [Drake, 1974: Chapters 4 & 6]. We say that κ is measurable iff there is some family U of subsets of κ, which satisfies the technical condition of being a κ-complete non-principal ultrafilter on κ. Very loosely, we can think of this as a generalisation of the ability to prove the *compactness theorem* in first order logic. A measurable cardinal κ allows us to prove the compactness of a logic which permits conjunctions and disjunctions of length $< \kappa$.[43] ω clearly enjoys this property since we can prove compactness for the language with conjunctions of length $< \omega$; i.e., first-order logic. It turns out we need need to go a long way, so to speak, before this property emerges again.

Measurable cardinals can influence the world of real numbers via determinacy theorems. We can describe this quite easily. Let us call a countably infinite sequence of natural numbers a *real number.*[44] Let A be a set of real numbers. A game on A consists of players I and II playing, in turns, natural numbers.

I	4		899		...
II		6		...	

After infinitely many turns of the game have been played out, player I wins if the resultant sequence $\langle 4, 6, 899, ... \rangle$ of natural numbers is in A and player II wins if not. The game on A is *determined* if there is a strategy that would allow one of the players to win. If determinacy holds of all sets of a particular kind, then we are able to establish a number of pleasing mathematical properties about that kind. For example, if games on sets of a particular kind are all determined, then the continuum hypothesis holds of those sets [Kanamori, 2003].[45]

The following theorem tells us that in the presence of a measurable cardinal, all the games of a certain natural type are determined.

Theorem 1. *(Martin) If there is a measurable cardinal, then all the games on sets of reals which are definable using just one just one quantification over the reals are determined.*[46]

---

[42]Strictly, I should mention the *sharps* here. However, these are not large cardinals but very special sets of natural numbers. While distracting, this point is still worth noting: see [Jech, 2003: Chapter 18]. For a possible counterexample to the claim above, we might consider Solovay's proof that given an inaccessible cardinal, there is a forcing extension $V[G]$ of the universe V that contains an inner model M in which, for example, every set of reals enjoys the perfect set property [Kanamori, 2003: Chapter 11]. There are a couple of things to say here. First, among the smaller large cardinals, inaccessibility is a particularly natural combinatorial property. Thus, it is perhaps unsurprising that it can have concrete (i.e., not merely logical) effects. But second, we should note that while the result above is stated in a model theoretic fashion, the use of forcing to get the inner model makes this a relative consistency result - not a straightforwardly combinatorial one.

[43]This is actually weaker or stronger than measurability depending on whether we show weak or strong compactness for the infinitary language. A notion of *medium compactness* fits measurability correctly [Chang and Keisler, 1973].

[44]A set of sequences of natural numbers is colloquially known as a set of *logician's real numbers*. The space of such sequences is homeomorphic with the irrational numbers: hence the relationship with the reals [Moschovakis, 1980].

[45]It should also be noted that if determinacy held for every sets of sequences of natural numbers A, then major distortions to set theoretic foundations would occur: e.g., the axiom of choice would fail to be true.

[46]Sets of reals definable in this way are usually known as analytic or coanalytic and are denoted as Σ^1_1 or Π^1_1 respectively.[Jech, 2003].

In this way, the augmentation of ZFC with large, large cardinals allows us to solve new problems about the familiar space of real numbers. For this reason, they are of *obvious* mathematical interest.[47]

This brings us to the question of what a superset theory offers set theoretic research. The answer is: not much. When we augment a theory of sets, like ZFC, with a theory of supersets, we do not get a significant increase in this kind of strength, even if we demand that the models of these theories are natural. While large cardinals are useful for finding natural models for superset theories, cardinals required for these purposes are not particularly large.[48] Indeed it seems unlikely that we would get anywhere near a measurable cardinal - our first example of an interesting large cardinal. The following phenomenon emerges:

Phenomena: For any *reasonable* superset theory Γ based on ZFC there is a natural model $\mathscr{M}$, which is simply derived from V_κ where $\kappa < \lambda$ the first Mahlo cardinal and $\mathscr{M} \models \Gamma$.

The first Mahlo cardinal is the least cardinal λ such that for every normal function there is a fixed point which is a regular, strong limit.[49] The idea behind the thought is that each of the examples of superset theories we have considered corresponds to the regular fixed point of some normal function. However, the thought is still somewhat vague given that our characterisation of a superset theory is not mathematically precise. A Mahlo cardinal is often thought of as the limit on the large cardinal axioms that can be defined, so to speak, from below [Drake, 1974]. Analogously, we build superset theories from below by cleverly concatenating chains of membership relations. A measurable cardinal, on the other hand, is defined, so to speak, from above by analogy with a property that is possessed by ω. As such, a measurable cardinal is much larger than a Mahlo cardinal. It is also more useful.[50]

The upshot of this discussion is to illustrate that contemporary set theory makes use of axioms that can construct models for any superset theory with trivial ease. Thus, if our goal is to solve problems about infinite collections, we should use ordinary set theory augmented with large, large cardinal axioms and not bother with superset theory at all. The superset approach is not only dispensable, it is mathematically uninteresting. The mathematician wants to formulate and solve precise problems and supersets are of negligible value for this enterprise. On the other hand, the philosopher is interested in a foundation for mathematics and must not naïvely shirk away from those areas which this foundation does not support. This does not mean that the philosopher must

[47] These results can be extended with further large cardinals using the work of Martin and Steel [Martin and Steel, 1989]. For example, in the presence of a even larger large cardinals, we may show that all the games on sets of sequences definable in the theory of the real numbers are determined.

[48] We should note that not all axiomatisations of superset theories imply the existence of large cardinals. For example, MK set theory (a set theory for a single layer of proper classes beyond ordinary sets) does not imply the existence of an inaccessible cardinal. Within MK we are able to prove that ZFC has a model, but the model could be unnatural. However, if we add an axiom to MK stating that ZFC has a natural model, then the existence of an inaccessible cardinal follows. Similar remarks apply to the theories $\Gamma_{SETS+SETS}$ and $\Gamma_{SETS\times SETS}$ discussed in the Appendix.

[49] A function $f : \lambda \to \lambda$ is *normal* if: it is *increasing* in the sense that for $\alpha < \beta$, $f(\alpha) < f(\beta)$; and *continuous* in the sense that for limit ordinals β, $f(\beta) = \bigcup_{\alpha<\beta} f(\alpha)$.

[50] Of course, we may want to countenance stronger superset theories than those based on merely ZFC. For example, we may decide the existence of a measurable cardinal is true and thus, should be an axiom of our system. If we then form a superset theory on this basis, Principle 2 would demand that there must be a measurable cardinal for every notion of membership utilised by the superset theory. In this case, a single measurable cardinal located in the lowest universe would suffice. After that, analogous natural models to those in the previous discussion will work for the ensuing universes and their membership relations. However, no signficant increase in strength will be gained over the original measurable cardinal assumption. This is, however, a relatively simply large cardinal assumption. It would be interesting to get a clearer picture of what would happen if, say, a proper class of supercompact or strong cardinals was demaned. Nonetheless, the phenomena above illustrates that the kind of transcendence given by these powerful large cardinals will not be made provided by natural models of superset theories.

abandon set theory or even its claim to providing a foundation for mathematics; merely, that this claim deserves reassessment.

4. Conclusion

The superset approach to membership provides an intellectually honest account of where the standard classical tools of set theory lead us when we come to ask what kind of thing the collection of all sets is. It takes us from the familiar world of proper classes into Grothendieck universes and beyond, for no matter how far we go, there will always be a further membership concept as yet unconsidered.

In comparison to other solutions on the market, the superset proposal could seem a little disappointing: we don't revise our logical resources; we don't appeal to novel metaphysical posits; and we don't just stipulate the problem away. Rather we simply apply the standard tools of classical logic and set theory while taking seriously the idea that something will always have been missed. It is a philosophical position rather than a mathematical one in that we cannot wrap up the content of the view in a simple first order theory. Indeed, this apparent deficiency comprises the deeper content of the view: indefinite extensibility is not just witnessed by an ever-escaping ontology, it is also brought to life in the essential inability to provide a total theory of membership.

References

JC Beall. *Spandrels of Truth*. Oxford University Press, 2009.

Andreas Blass. Are proper classes objects?, 2011. URL `http://mathoverflow.net/questions/71765/are-proper-classes-objects/71773#71773`.

Andreas Blass. The kunen inconsistency and definable classes, 2013. URL `http://mathoverflow.net/questions/129498/the-kunen-inconsistency-and-definable-classes`.

George Boolos. To be is to be the value of a variable (or to be some values of some variables). *The Journal of Philosophy*, 81:430–450, 1984.

George Boolos. Must we believe in set theory? In *Logic, Logic, and Logic*, pages 120–132. Harvard University Press, 1998a.

George Boolos. Reply to charles parsons' 'sets and classes'. In *Logic, Logic, and Logic*, pages 120–132. Harvard University Press, 1998b.

Nicolas Bourbaki. Univers. In Michael Artin, Alexandre Grothendieck, and Jean-Louis Verdier, editors, *Séminaire de Géométrie Algébrique du Bois Marie - 1963-64 - Théorie des topos et cohomologie étale des schémas - (SGA 4) - vol. 1 (Lecture notes in mathematics 269)*. Springer-Verlag, Berlin, 1972.

C. C. Chang and H. J. Keisler. *Model Theory*. North Holland Publishing Company, 1973.

Alonzo Church. A formulation of the simple theory of types. *Journal of Symbolic Logic*, 5(2):55–68, 1940.

Paul J. Cohen. *Set Theory and the Continuum Hypothesis*. W. A. Benjamin Inc., New York, 1966.

Roy T. Cook. Embracing revenge: On the indefinite extendibility of language. In J. C. Beall, editor, *Revenge of the Liar: New Essays on the Paradox*, page 31. Oxford University Press, 2007.

F.R. Drake. *Set Theory: An Introduction to Large Cardinals*. North-Holland, 1974.

Michael A. E. Dummett. *The Seas of Language*, volume 58. Oxford University Press, 1993.

Solomon Feferman and G. Kreisel. Set-theoretical foundations of category theory. In *Reports of the Midwest Category Seminar III*, volume 106 of *Lecture Notes in Mathematics*, pages 201–247. Springer Berlin Heidelberg, 1969. ISBN 978-3-540-04625-7. doi: 10.1007/BFb0059148. URL `http://dx.doi.org/10.1007/BFb0059148`.

Hartry Field. *Saving Truth from Paradox*. OUP, Oxford, 2008.

Kit Fine. Our knowledge of mathematical objects. In Tamar Szabo Gendler and John Hawthorne, editors, *Oxford Studies in Epistemology Volume 1*, page 89. Oup Oxford, 2005.

Thomas Forster. The iterative conception of set. *Review of Symbolic Logic*, 1(1):97–110, 2008.

Michael Glanzberg. A contextual-hierarchical approach to truth and the liar paradox. *Journal of Philosophical Logic*, 33(1):27–88, 2004.

Michael Glanzberg. Context and unrestricted quantification. In *Absolute Generality*, pages 45–74. Oxford University Press, 2006.

Volker Halbach. Tarski hierarchies. *Erkenntnis*, 43:339–367, 1995.

Volker Halbach. *Axiomatic Theories of Truth*. Cambridge University Press, London, 2011.

P.R. Halmos. *Naive Set Theory*. Undergraduate Texts in Mathematics. Springer, 1960.

Allen Hazen. Against pluralism. *Australasian Journal of Philosophy*, 71(2):132–134, 1993.

Geoffrey Hellman. *Mathematics Without Numbers: Towards a Modal-Structural Interpretation*. Oxford University Press, 1989.

Thomas Jech. *Set Theory*. Springer, Heidelberg, 2003.

A. Kanamori. *The Higher Infinite: Large Cardinals in Set Theory from Their Beginnings*. Springer, 2003.

Saul Kripke. Outline of a theory of truth. *Journal of Philosophy*, 72:690–716, 1975.

Kenneth Kunen. *Set Theory: an introduction to independence proofs*. Elsevier, Sydney, 2006.

F. W. Lawvere. Diagonal arguments and cartesian closed categories. In P. Hilton, editor, *Category Theory, Homology Theory and their Applications II*, pages 134–145. Springer, 1969.

Hannes Leitgeb. What truth depends on. *The Journal of Philosophical Logic*, 34:155–192, 2005.

Hannes Leitgeb. What truth should be like (but cannot be). *Philosophical Compass*, 2(2):276–290, 2007.

Øystein Linnebo. Pluralities and sets. *Journal of Philosophy*, 107(3), 2010.

Øystein Linnebo. Reference by abstraction. *Proceedings of the Aristotelian Society*, 112(1pt1):45–71, 2012a.

Øystein Linnebo. Grounded properties. 2012b.

Donald A. Martin. Multiple universes of sets and indeterminate truth values. *Topoi*, 20(1):5–16, 2001.

Donald A. Martin and John R. Steel. A proof of projective determinacy. *Journal of the American Mathematical Society*, 2(1):pp. 71–125, 1989.

John Mayberry. On the consistency problem for set theory: An essay on the cantorian foundations of classical mathematics (i). *British Journal for the Philosophy of Science*, 28(1):1–34, 1977.

Colin McLarty. What does it take to prove fermat's last theorem? grothendieck and the logic of number theory. *Bulletin of Symbolic Logic*, 16(3):359–377, 09 2010. doi: 10.2178/bsl/1286284558.

Y.N. Moschovakis. *Descriptive Set Theory*. North Holland, 1980.

Edward Nelson. *Predicative Arithmetic*. Princeton University Press, 1986.

Charles Parsons. *Mathematics in Philosophy: selected essays*. Cornell University Press, New York, 1983.

Graham Priest. The logic of paradox. *Journal of Philosophical Logic*, 8:219–241, 1979.

Graham Priest. The structure of the paradoxes of self-reference. *Mind*, 103(409):25–34, 1994.

Graham Priest. *Beyond the Limits of Thought*. Oxford University Press, 2002.

W. N. Reinhardt. Set existence principles of Shoenfield, Ackermann and Powell. *Fundamenta Mathematicae*, 84(1):5–34, 1974.

Michael Resnik. Second-order logic still wild. *The Journal of Philosophy*, 85(2):75–87, 1988.

Philippe Schlenker. Super liars. *Review of Symbolic Logic*, 3(3):374–414, 2010.

Stewart Shapiro. *Foundations without Foundationalism: a case for second order logic*. OUP, Oxford, 1991.

Stewart Shapiro and Crispin Wright. All things indefinitely extensible. In Agustin Rayo and Gabriel Uzquiano, editors, *Absolute Generality*. Oxford University Press, London, 2007.

J. P. Studd. The iterative conception of set: A (bi-)modal axiomatisation. *Journal of Philosophical Logic*, pages 697–725, 2012.

Alfred Tarski. The concept of truth in formalized languages. In John Corcoran, editor, *Logic, semantics, metamathematics: papers from 1923 to 1938*. Clarendon Press, Oxford, 1956.

Gabriel Uzquiano. Varieties of indefinite extensibility. *Notre Dame Journal of Formal Logic*, forthcoming.

Hao Wang. Truth definitions and consistency proofs. *Transactions of the American Mathematical Society*, 73 (2):243–275, 1952.

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

Zach Weber. Extensionality and restriction in naive set theory. *Studia Logica*, 94(1), 2010b.

Timothy Williamson. Indefinite extensibility. *Grazer Philosophische Studien*, 55:1–24, 1998.

Timothy Williamson. Everything. *Philosophical Perspectives*, 17(1):415–8211, 2003.

Ernst Zermelo. On boundary numbers and domains of sets: new investigations in the foundations of set theory. In *From Kant to Hilbert: A Source Book in the Foundations of Mathematics*. Oxford University Press, 1976.

Appendix: Axiomatising superset theories

In this appendix, we state axiomatisations and some basic results for $SETS + SETS$ and $SETS \times SETS$.

An axiomatisation of $SETS + SETS$. Recall that we want to build one notion of membership η on top of another $\in$. As we've discussed earlier, it is probably easier to visualise what we are doing by thinking of $\in$ as being about a lower universe. Moreover, we could have introduced a constant symbol U, in the manner of [Feferman and Kreisel, 1969] or [Bourbaki, 1972] and produced an axiomatisation from there. This would have worked and, indeed, we'll exploit this relationship below; however, we opted against this in our official axiomatisation in order to emphasise that the ambiguity of supersets is located in notion of membership itself.

Let $\mathscr{L}_{\in,\eta} = \{\in, \eta\}$.

Definition 2. Let us say that x is an *$\in$-set* if

$$\exists y\, x \in y.$$

Otherwise, we shall say that x is a *proper η-class*.

Remark 3. The idea here is that x is an $\in$-set if it can be, so to speak, covered using the $\in$ relation. We'll place further constraints on $\in$-sets in the axioms below.

Write $\forall^{\in} x\, \varphi(x)$ in place of $\forall x(x \text{ is an } \in\text{-set} \to \varphi(x))$.

Let $ZFC(\in)$ be the result of writing out the axioms of ZFC and using $\forall^{\in}$ instead of $\forall$.

Let $ZFC(\eta)$ be the result of writing out the axiom of ZFC with η replacing $\in$. If t is any (defined) term of ordinary set theory, let t^{η} be the result of defining it with η instead of $\in$.

Let our theory $\Gamma_{SETS+SETS}$ consist of the following axioms:

(1) $ZFC(\in)$.
(2) $ZFC(\eta)$.
(3) $\forall x \forall y (x \in y \rightarrow x \eta y)$ (*Cumulativity*).
(4) $\forall x \forall y (x \in y \wedge x$ is an $\in$-set $\rightarrow y$ is an $\in$-set) (*End-extension*).[51]
(5) $\forall x \forall y (rank^{\eta}(x) \leq^{\eta} rank^{\eta}(y) \wedge y$ is an $\in$-set $\rightarrow x$ is an $\in$-set) (*Top extension*).[52]
(6) $\exists x (\forall^{\in} y \; y \eta x)$ (*Closure*).

Lemma 4. *(i) There is an r such that for all $\in$-sets y*

$$y \eta r \leftrightarrow y \notin y.$$

Proof. By (6), let u, be such that $\forall^{\in} y \; y \eta u$. Using η-separation, we see that there is some r such that[53]

$$\forall y (y \eta r \leftrightarrow y \eta u \wedge \neg y \eta y).$$

□

Theorem 5. *Suppose there are two inaccessible cardinals $\kappa_1 < \kappa_2$. Let $\mathscr{M} = \langle V_{\kappa_2}, \in \restriction (V_{\kappa_1} \times V_{\kappa_1}), \in \restriction (V_{\kappa_2} \times V_{\kappa_2}) \rangle$. Then $\mathscr{M}$ is a natural model and*

$$\mathscr{M} \models \Gamma_{SETS+SETS}.$$

In fact, if we forgo the restriction to natural models, consistency may be established more easily via an adaptation of a trick from Feferman and Cohen [1969, 1966].

Theorem 6. $Con(ZFC) \rightarrow Con(\Gamma_{SETS+SETS})$.

Proof. To prove this we first observe that $\Gamma_{SETS+SETS}$ is mutually interpretable with the theory, Δ, articulated in the a language $\mathscr{L} = \{\in, U\}$ which admits a constant for universes, and consisting of the following axioms:

(1) ZFC;
(2) $ZFC(U)$ - where we restrict all of the quantifiers in axioms to U; and
(3) $\exists \alpha \; U = V_{\alpha}$.

Thus, it suffices to show that $Con(ZFC) \rightarrow Con(\Delta)$. Suppose not, then $Con(ZFC)$ and for some finite $\Lambda \subseteq \Delta$, Λ is unsatisfiable. Let $\Lambda = \Lambda_0 \cup \Lambda_1 \cup \{\exists \alpha \; U = V_{\alpha}\}$ where $\Lambda_0 \subseteq ZFC$ and $\Lambda_1 \subseteq ZFC(U)$. Since $Con(ZFC)$ we may fix some $\mathscr{M} \models ZFC$. Clearly, $\mathscr{M} \not\models \Lambda$.

Let $\Lambda_1^{\dagger}$ be the result of substituting ordinary quantifiers for the restricted quantifiers in Λ_1. Then, using reflection we know that $ZFC \vdash$ "$\exists \alpha \; V_{\alpha} \models \Lambda_1^{\dagger}$". Thus, it can be easily seen that $\mathscr{M} \models \Lambda$: contradiction. □

[51] This axiom isn't strictly necessary as it's implied by (5), however, in the absence of (5) it could provide a pleasing weakening of the system.
[52] In ordinary set theory, the rank function takes a set x to the least α such that $x \in V_{\alpha+1}$. This can be represented by a term, *rank*, whose definition can then be relativised to obtain $rank^{\eta}$ by replacing all instances of $\in$ by η.
[53] Of course this follows more directly from (6) and $\in$-Foundation.

An axiomatisation of $SETS \times SETS$. Recall that we are trying to stack indefinitely many membership relations on top of each other and that we are indexing our membership relation to do this.

Let $\mathscr{L} = \{\in\}$ where $\in$ is now a three-place predicate, where the third argument is written as a subscript on $\in$.

Definition 7. (i) For all x, y

$$x \ll y \leftrightarrow \exists z\, x \in_z y.$$

This gives us a notion of universal membership: membership according to any relativisation.

(ii) x is a *super ordinal*, abbreviated $SupOn(y)$, if there is some y such that x is $\in_y$-transitive and an $\in_y$-linear order. We are going to use the super ordinals to index the membership relation.

(iii) If x is a super ordinal, we say that y is an $\in_x$*-set* if

$$\exists z\, y \in_x z.$$

The idea here is that y is an $\in_x$-set if it can, so to speak, be covered using the $\in_x$ relation.

Let us write $\forall^y z\, \varphi(z)$ for

$$\forall z (z \text{ is a } \in_y \text{-set } \to \varphi(z)).$$

Let $\forall y(SupOn(y) \to ZFC(\in_y))$ be the result taking each axiom of ZFC:

- adding the subscript $_y$ to each $\in$;
- changing quantifiers to $\forall^y$; and
- substituting that into the space in $\forall y(SupOn(y) \to ...)$.

Remark. We should note that $\forall y(SupOn(y) \to ZFC(\in_y))$ is not a single sentence but an infinite collection of sentences.

Let $ZFC(\ll)$ be the result of replacing $\in$ by $\ll$ in every axiom of ZFC. Let $\Gamma_{SETS\times SETS}$ comprise of the following axioms:

(1) If $x \in_z y$, then z is a super-ordinal.
(2) There is some x such that for all y and for all z, $x \notin_y z$, which we denote $\emptyset$.
(3) $\emptyset$ is a super-ordinal.
(4) $\forall y(SupOn(y) \to ZFC(\in_y)$ for all y.
(5) $ZFC(\ll)$.
(6) If x and y are super-ordinals, and $x \ll y$, then $\forall u \forall w(u \in_x w \to u \in_y w)$ (*Cumulativity*).
(7) If $rank^{\ll}(x) \leq^{\ll} rank^{\ll}(y)$ and y is a $\in_z$-set, then x is a $\in_z$-set (*Top extension*).
(8) If x is a super-ordinal, then there is some super-ordinal $y \gg x$ and some u such that $\forall_x z\, z \in_y u$; moreover, for every $y \gg x$ there is such a u (*Closure*).

Proposition 8. *(i) Any ordinary ordinal (i.e., $\in_\emptyset$-ordinal) α is a super ordinal.*

(ii) For all super-ordinals x, there is a Russell's set.

Theorem 9. *Let κ be an inaccessible limit of inaccessible cardinals and let $\langle \lambda_\alpha | \alpha < \kappa \rangle$ enumerate all the inaccessible cardinals below κ. Then*

$$\langle V_\kappa, E \rangle \models \Gamma_{SETS\times SETS}$$

where

$$E = \{(x,y,\alpha) \in V_\kappa \times V_\kappa \times On^{V_\kappa} \mid \alpha \textit{ is inaccessible } \wedge x \in y \wedge rank(y) < \lambda_\alpha\}.$$

Moreover, this is the smallest natural model of $\Gamma_{SETS\times SETS}$.

Again, if we forgo the natural models requirement, the theory is no stronger than ZFC.

Theorem 10. $Con(ZFC) \to Con(\Gamma_{SETS\times SETS})$.

Proof. As with $\Gamma_{SETS+SETS}$, we first note that $\Gamma_{SETS\times SETS}$ is mutually interpretable with a theory that it a little easier to work with. Let Δ be a theory articulated in $\mathscr{L} = \{\in, u\}$ where u is a one-place function symbol. Let Δ consist of the following sentences:

(1) ZFC;
(2) $\forall\alpha\ \varphi^{u(\alpha)}$ where φ is an axiom of ZFC;
(3) $\forall\alpha\exists\kappa\ u(\alpha) = V_\kappa$.
(4) $\forall\alpha < \beta\ u(\alpha) \subsetneq u(\beta)$.

It then suffices to show that $Con(ZFC) \to Con(\Delta)$. Suppose not. Then $Con(ZFC)$ and there is some finite $\Lambda \subseteq \Delta$ such that Λ is unsatisfiable. Let $\Lambda = \Lambda_0 \cup \Lambda_1 \cup \{\forall\alpha\exists\kappa\ u(\alpha) = V_\kappa, \forall\alpha < \beta\ u(\alpha) \subsetneq u(\beta)\}$ where $\Lambda_0 \subseteq ZFC$ and $\Lambda \subseteq \{\forall\alpha\varphi^{u(\alpha)} \mid \alpha \in On \wedge \varphi \in ZFC\}$. Using our assumption, fix some $\mathscr{M}$ such that $\mathscr{M} \models ZFC$. Then clearly $\mathscr{M} \not\models \Lambda$.

Let $\Lambda_1^\dagger$ be the result removing the initial $\forall\alpha$ and replacing the restricted quantifiers by ordinary quantifiers for sentences in Λ_1. By reflection, we know that $ZFC \vdash$ "$\forall\alpha\exists\beta > \alpha\ V_\alpha \models \Lambda_1^\dagger$". Use this to fix the interpretation of u function. Then from here it can be shown that $\mathscr{M} \models \Lambda$: contradiction. □

UNIVERSITY OF ABERDEEN