Chapter 1

Introduction

"Admittedly, the present state of affairs where we run up against the paradoxes is intolerable. Just think, the definitions and deductive methods which everyone learns, teaches, and uses in mathematics, the paragon of truth and certitude, lead to absurdities! If mathematical thinking is defective, where are we to find truth and certitude?[1]*"*

– David Hilbert, *On the Infinite*

[1] translation from [3]

1.1 Mathematics as a Touchstone and the Centrality of Set Theory

Mathematical proofs provide a touchstone of clarity and convincingness which serves as an inspiration to philosophy and other disciplines. While it is possible to doubt the results of mainstream mathematical arguments (philosophers are capable of doubting anything), there's something striking about just how convincing arguments mathematical proofs are. Consider the standard argument that there are infinitely many primes. Even philosophers who deny that there are numbers (and hence think the argument as usually stated is unsound) are strongly tempted to say that we know *something like the premises* and that the proofs provide some kind of valuable amplification of this knowledge. The premises we use in informal mathematical reasoning have a combination of prima facie obviousness and power/generality, which makes them exemplary tools for expanding our knowledge and resolving disputes in cases where people's initial hunches disagree. It's no surprise that Leibniz[2] wished philosophers could resolve their disputes like mathematicians by saying 'let us calculate' (or at least, 'let us each look for a proof').

Now (in many ways) set theory lies at the heart of modern mathematics, and it does powerful mathematical (not just philosophical) work as a foundation for the whole. So one might hope that the set theoretic foundations for mathematics would share the clarity and convincingness we hope for from

[2] See page 14 of [17].

mathematical arguments.

But certain problems in the philosophical foundations of set theory raise serious questions and doubts about our acceptance of the axioms of set theory. These questions are more subject matter specific and threatening to things normal mathematicians care about than generic philosophical doubts about whether there are any abstract objects, or whether the knowledge we get from the standard proof that there are infinitely many primes is better construed nominalistically or platonistically.

Specifically, the development of set theory resolved a great many problems in analysis. And it provided a formal framework to allow interactions between various areas of mathematics creating, as Hilbert famously observed [47], a kind of mathematical paradise. However contradiction threatened Hilbert's paradise, in the form of Russell's paradox. This problem was practically solved by accepting the iterative hierarchy conception of sets and the standard Zermelo-Fraenkel with choice (ZFC) first order axioms for set theory. And now set theory is widely accepted as a foundation for all of modern mathematics. It is hard to deny that the mathematical results which are currently stated in terms of set theory reflect genuine and important knowledge of some kind. But a question of **how to justify these axioms** remains.

So we may ask: is the price of remaining in Cantor's paradise giving up the old ambition of founding mathematics on intrinsically obvious seeming principles?

In this book I will develop a unified determinate conception of set theoretic truth, which vindicates many of our intuitive expectations regarding set theory. My proposal attempts to improve on standard 'actualist' approaches to set theory, which fall short of this ideal in several ways.

1.2 Actualism and its Discontents

According to standard actualist approaches to set theory, set theory is about objects called 'sets', which exist outside of space and time. On this view, sets are abstract mathematical objects, just like the natural numbers (on a platonist understanding of the natural numbers). Apparent existence claims made by set theorists (like 'there is a set which has no elements') are made true by the existence of corresponding objects, just like ordinary existence claims about cities or electrons or cars.

Actualists run into three problems. First, actualist approaches don't offer a determinate conception of set theoretic structure. In particular, the height of the hierarchy of sets is left vague or mysterious. As the Burali-Forti paradox[3] dramatizes, it appears that, for any height that the hierarchy of sets could achieve, it would be possible to have a strictly larger structure which adds an extra layer of 'sets' on top of it. So it seems arbitrary to suppose that the hierarchy of sets just stops somewhere.

Second, as a foundation for mathematics, one might hope that set theory should be able to represent any mathematical structure one might want to

[3] See section 2.2 for more details about this.

study. And the idea that set theory has this kind of generality is prima facie quite intuitive. But actualist set theory throws up its hands for structures that are 'too large.' So actualism makes it hard to capture the intuition that 'any possible structure' should, in some sense, be fair game for treatment within set theory.

Third, there's a problem about intuitively justifying certain axioms of set theory. We normally hope that axioms of mathematics, which are taken as starting points, will be extremely prima facie plausible (if not completely indubitable or impossible to empirically cast doubt upon). So one might hope that, once we adopt a good philosophy of set theory, all the ZFC axioms (the most widely used axioms for set theory) will seem clearly true, or at least justifiable via principles that are obvious (or obvious relative to the iterative hierarchy conception of sets and the presumption that some abstract platonic structure corresponding to that structure exits).

However, philosophers have had significant difficulty in finding any such justification for certain of the ZFC axioms of set theory. For example, in [82], Hilary Putnam writes "Quite frankly, I see no intuitive basis at all for . . . the axiom of replacement. Better put, I do not see that a notion of set on which that axiom is clearly true has ever been explained." Instead philosophers of mathematics and mathematicians have made do with more external justifications, via things like the fact that in a century of working with the ZFC axioms mathematicians haven't found a contradiction, or via the fact that contested axioms like replacement seem to be fruitful in plausible consequences and can speed up the proofs of things that we have antecedent mathematical reason for believing. But this state of affairs can feel unsatisfying.

1.3 A New Flavor of Potentialism

In this book I will advocate and develop a different way of understanding set theory, and then use it to provide a justification for the ZFC axioms which avoids the three problems above.

In Part I I will argue that we should reject actualism about set theory in favor of an idea called set theoretic potentialism, which has been developed by philosophers like Putnam, Parsons, Hellman and Linnebo. The key idea behind potentialism is that, rather than taking there to be a hierarchy of sets which stops at some particular point, we can think about set theorists as making modal claims about what hierarchy-of-sets-like structures are possible and how such structures could possibly be extended.

Merely accepting potentialism solves one of the problems for actualism above. For we are not committed to postulating an arbitrary or vague height for the hierarchy of sets. And it arguably puts us in a better position to honor the intuition that 'any possible structure' can be studied within set theory. However, existing forms of potentialist set theory don't tend to (even claim to) make progress on the justification problem above. Moreover, I will argue that each faces a number of further philosophical problems. For example, Putnam and Parsons underspecify which modal notion to use in formulating potentialist set theory. Linnebo faces difficulties cashing out his preferred notion of possibility without committing himself to a linguistically weird double role for set talk and/or reinventing actualism's problem re: commitment to an arbitrary stopping point to the hierarchy of sets. And Hellman is explicitly conflicted about his version of potentialism's commitment to second order logical comprehension principles and/or mereology.

In Part II I will develop a slightly different flavor of potentialism which avoids these problems for existing formulations. In a nutshell, I will propose that set theoretic statements are best understood as claims about a kind of logical possibility and extendability which constrains the behavior of all objects.

In Part III I will then tackle the core remaining problem: justifying the ZFC axioms. I'll offer a formal system for reasoning about logical possibility whose principles are intrinsically appealing, and prima facie obvious in the way that we traditionally expect mathematical axioms to be. And I will show that reasoning in this formal system suffices to justify acceptance of (potentialist translations of) all the axioms of ZFC set theory. This allows us to resolve the problem above by providing an internal justification for replacement (and all the other standard set theoretic axioms).

Finally in Part IV, I will consider (in a more speculative vein) how the potentialist approach to set theory advocated above can be fit into a unified larger philosophy of mathematics.

1.4 Outline

My plan of action will be as follows.

In chapter 2 I will introduce the standard ‘actualist’ approach to set theory, and note how it faces an arbitrariness problem (which is highlighted by the Burali-Forti paradox), as well as a problem about justifying our use of certain axioms of potentialist set theory. I will review some of the main actualist approaches to solving this problem, and argue that they are unsatisfactory.

In chapter 3 I will introduce a ‘potentialist’ approach to set theory which was sketched by Hillary Putnam in [83] and has been developed in varying ways by Parsons, Linnebo and Hellman since. I will note that adopting this potentialist approach promises to let us avoid the arbitrariness woes which beset actualist set theory. However, serious philosophical questions and disagreements arise when we consider how to understand the notions of possibility and extendability which potentialist approaches invoke. Also the potentialist is still on the hook to show that they can show that their approach is not only compatable with the usual ZFC axiomitization of set theory but that these axioms can be justified from a simple unified potentialist conception of the set theoretic hierarchy.

In chapter 6 I will argue that independent considerations in the philosophy of logic motivate accepting a suitably ‘free standing’ notion of logical possibility, which is not itself defined in terms of set theory. Then I’ll suggest that this notion usefully generalizes to a corresponding notion of logically possible extendability.

In chapters 7 and 8 I will sketch how we can use this notion of logical possibility to streamline formulate potentialist set theory. And in 9 I will argue that doing so avoids the problems facing existing versions of potentialist set theory discussed above. In chapter 10 I will argue that my case for potentialist set theory should be taken seriously even by philosophers with strong naturalist inclinations (despite a worry suggested by Burgess and Rosen's famous dilemma for mathematical nominalists in [14]).

I will then turn to the justification problem noted above. In chapters 12 through 14 I will introduce some general and intuitive methods of reasoning about logical possibility. And in chapters 15 and 16 I will show that, happily, adopting my preferred potentialist set theory lets us reconstruct all the standard ZFC reasoning about set theory using axioms and methods of inference which are just as prima facie intuitive and compelling as we normally expect mathematical axioms to be. Thus the familiar hope that mathematical proofs can be justified on the basis of principles that seem prima facie obvious (if not completely indubitable) can be maintained if we accept the account of the nature of set theory which I propose.

In final section of the book, I'll consider whether and how the picture of set theory I have advocated can be fit into an attractive larger philosophy of mathematics. In chapter17 I'll show how my potentialism about set theory can be extended in a natural and unified way to a general nominalism about mathematical objects.

In chapters 18 through 22 I will consider an important and influential challenge to mathematical nomianlism: the Quinean Indispensability argument.

Since, my potentialistic approach requires nominalism about the sets some (though I do not) may be persuaded that defending potentialism requires a broader defense of nominalism. To that end, I will argue for cautious optimism about defending the nominalism of chapter 17 from this objection. In particular, using a framework I develop in chapter 19 I'll argue that the arguments made suggesting the nominalist is in a worse position than the platonist are either facially unconvincing or can be positively countered by transforming a platonistic paraphrase of our scientific theories into a nominalistic one.

In chapter 23 I'll lay out my preferred deflationary realist (very broadly neo-carnapian) approach to mathematical objects outside higher set theory, as an alternative to the nominalism above. I'll note that although this view avoids the classic Quinean Indispensibility argument (by accepting the existence of mathematical objects), a different version of the indispensibility argument can seem to arise for it. Then I will argue that this problem can be solved.

In chapter 24 I'll note that both approaches to general mathematics just mentioned (i.e., the nominalist approach and the deflationary realist one) accord with some traditionally popular logicist and structuralist ideas about the nature of mathematics.

Finally, in chapter 25 I will conclude with some even more brief and speculative claims about the relationship between the approach to set theory I've defended here and some other traditional big questions in the philosophy of mathematics.

1.5 Relation To Mathematical and Philosophical Practice

Let me finish this introduction with some quick caveats about the nature and aim of my project.

First, I don't claim set theorists should literally rewrite textbooks and journal articles in potentialist terms. Mathematicians' current practice of (making arguments which can be reconstructed as) proving things from in first order logic from the ZFC axioms is fine. And doing something like logical deduction from purely first order axioms may be unavoidably easier (for minds like ours) than thinking about elaborate modal extendability claims. If one thinks about apparent first order claims in mathematics as abbreviating claims about potentialist claims, then the main result in this book shows that it's unnecessary to unpack this abbreviation in mathematical contexts (because the ZFC axioms and everything derivable from them must also be true on a potentialist reading).

However, there's a sense in which I *am* suggesting potentialist paraphrases are what people should mean when they do set theory – or at least when they think about set theory in philosophical contexts. They 'should' replace current set theory with the potentialist version of because understanding set theory potentialistically blocks various intuitive puzzles, and makes sense of things that we normally want to say about set theory.

One can think of my current project of developing potentialist foundations for set theory as analogous to the familiar project of providing a set theoretic foundation for analysis. Our naive reasoning about certain concepts (limits in one case, the height of the hierarchy of sets in the other) turned out to lead to paradox. So it is desirable to find a different way of thinking about the relevant mathematical concepts which will let us capture the intuitive mathematical significance and interest of relevant mathematical claims while blocking paradoxical inferences. And it is desirable to cash out old mathematical concepts, which paradoxes may have led us to doubt that we have a coherent grip on, in other terms which we seem to understand in a way that does not invite paradox. I argue that if we cash out standard set theory in potentialist terms, the Burali-Forti paradox does not arise and yet all of mathematicians' ordinary reasoning about set theory is justified.

Second, the potentialist understanding of pure set theory which I advocate is compatible with a range of different views about how to understand talk of other kinds of pure mathematical structures, like the natural numbers, (which we seem to have a definite conception of that does not give rise to a version of the Burali-Forti paradox and arbitrariness worries noted above)[4]. For example, it is compatible with the approach to set theory

[4]One *might* follow orthodox set theoretic foundationalism, and take claims about all such claims to be shorthand for statements in the language of set theory. If one takes this view, then adopting my potentialist approach to set theory will mean accepting general nominalism about mathematical objects (on which all apparent mathematical existence claims are really just a shorthand for corresponding modal claims).

However, there are also some motivations refusing to identify talk of *other* mathematical structures (like the natural numbers) with claims about set theory and this approach to mathematical objects is optional. I personally think it's attractive to follow Benacerraf's [2] suggestion that we should ceterus paribus treat quantification over such mathematical structures in the same way that quantification over cities and holes which it superficially resembles. And elsewhere I advocate a quantifier variance view on which existence claims about such objects are literally true, even if (like cities and holes and shadows) some such objects may not be metaphysically fundamental (e.g., one might think about the truth of existence claims about mathematical structures in different languages as being grounded

I will advocate to take the natural numbers to be free standing actualist objects. It also compatible with my approach to interpret claims about the natural numbers as set theoretic statements in disguise (a la Bourbaki), and plug them into the potentialist paraphrase for set theory I propose, yielding a more nominalist view.

Third, I will (largely) bracket questions about how applied mathematics is best understood/developed in this book. I think ultimately it's important that whatever one says about pure mathematics should harmonize with what one says about applied mathematics. And I will make some brief suggestions about how to handle applied mathematics (and claims about mathematical objects other than the sets) in the conclusion. However, my task here will merely be to highlight an attractive solution to some intrinsic problems about how to understand pure set theory and I leave solving philosophical problems about applied mathematics for another place and time.

Finally, I claim to be providing a 'logical' foundation for potentialist set theory. I take this foundation to rest entirely on intuitively compelling principles which are 'subject-matter neutral' and constrain the behavior of all objects (and thus perhaps accord with Frege's criterion for logicality). But I don't mean to claim that the foundational principles which I will suggest are analytic, cognitively trivial, or impossible for any rational being to doubt.

---

in facts about logical possibility).

But, for present purposes, I just want to note that I don't take the philosophy of set theory which I develop and advocate in this book to commit one to adopting a nominalist or potentialist understanding of other mathematical structures.

Chapter 2

Actualist Set Theory and Problems for it

2.1 Actualist Set Theory and The Iterative Hierarchy Conception

On a straightforward actualist approach to set theory, there are abstract objects called 'the sets', much as there are abstract objects called 'the natural numbers.' And we can ask: what sets exist? And what kind of structure do the sets have under the relation of membership?

Naively one might want to say that, for any formula $\phi(x)$, there is a set whose elements are exactly those objects that satisfy ϕ. But, as Bertrand Russell famously showed, this leads to paradox via the conclusion that there must be a set whose elements are exactly the sets which aren't members of themselves.

The (widely embraced) iterative hierarchy conception of the sets solves this problem by suggesting a different picture of what sets exist. On this picture we should think about the sets as forming layers, with sets at a given layer in the hierarchy only being able to have elements which are available at previous layers. Following Boolos[11] one can spell out this idea out by saying that the hierarchy of sets consists of a two sorted structure consisting of:

- a well ordered series of stages, with no last element, and
- a collection of sets formed at these stages, such that a set is formed at a stage iff its members are all formed at earlier stages.

One can think of the iterative hierarchy conception as partially specifying a structure for the (pure) sets. If we adopt the idea of a hierarchy of sets, then the principles above specify an intended width for this structure.

However, the principles above do not specify an intended height for the hierarchy of sets (since there are many different logically possible well orderings which do not have a last element, e.g., ω, $\omega + \omega$ etc.). And there are well-known reasons for doubting that we have any coherent and adequate conception of absolute infinity (the supposed height of the hierarchy of sets).

2.2 A Burali-Forti/Arbitrariness Problem

The concern here (i.e., the concern about our notion of absolute infinity and the height of the hierarchy of sets) is not simply that it might be impossible to cash the notion of absolute infinity out in other terms. After all, every theory will have to take some notions as primitive. Rather, philosophical worries arise from the appearance that it is logically impossible for any collection of objects to satisfy our intuitive notion of absolute infinity – just as Russell's paradox shows that it's logically impossible for any collection of objects to satisfy the axioms of naive set theory.

A common intuitive conception of the hierarchy of sets says that the hierarchy of sets goes 'all the way up' – so no restrictive ideas of where it stops are needed to understand its behavior. However, if the sets really do go 'all the way up' in this sense, then it would seem that they should satisfy the following well-ordering principle.

> For any way some things could be well-ordered, there is an ordinal corresponding to it.

But the ordinals themselves are well-ordered, and there is no ordinal corresponding to this well-ordering, i.e., there is no ordinal which has the same order-type as the class of all ordinals. Thus (it would seem), the naive well ordering principle above can't be correct. The simplest response to this problem might be to find some other restrictive characterization of the sets (in particular, some other characterization of the intended height of the hierarchy of sets)[1]. However, it's not clear that *any* intuitive conception of the intended height of the sets remains once the paradoxical well-ordering principle above is retracted. As Wright and Shapiro put it [97], all our reasons for thinking that sets exist in the first place appear to suggest that, for any given height which an actual mathematical structure could have, the sets should continue up past this height.

It seems arbitrary to say that the hierarchy of sets just stops somewhere, if a suitable stopping point is not pinned down by something in our conception of the hierarchy of sets[2]. Saying that the hierarchy of sets just happens to stop at a certain point seems to violate intuitive principles of metaphysical parsimony. For one seems committed to positing an extra - otherwise entirely unmotivated - joint in reality, namely the height of the hierarchy of sets. One might also worry about the epistemology of this stopping point, and why we should think set theorists' reasoning about large cardinals etc.

[1] Note that the axioms of ZFC and even ZFC_2 don't suffice to categorically determine the height.

[2] To clarify this worry, note that I'm not suggesting the actualist must think the hierarchy of sets 'must stop somewhere' in the sense that they must say there's a largest ordinal. There's no problem about saying that every ordinal has a successor (as indeed is required Boolos' version of the iterative hierarchy of sets). There's no problem about there saying that for every set/ordinal x there's a strictly larger set/ordinal y.

The problem is that the actualist takes there to be some plurality of objects (the sets) forming an iterative hierarchy structure. And (by the modal principle just mentioned) it seems that for any plurality of objects satisfying the conception of an iterative hierarchy above, it seems that it would be possible for there to be a strictly lager iterative hierarchy structure, which mirrors the original structure, but adds a new *limit ordinal* above all he ordinals within the original structure and a layer of classes corresponding to all possible ways of choosing some objects from the original plurality. And it seems that the resulting structure generated would answer everything in our conception of the hierarchy of sets just as well as the original structure did.

Thus the actualist seems forced to say that the plurality of existing sets just happens to instantiate one possible/logically coherent structure satisfying the iterative hierarchy conception of sets rather than another conception which satisfies this conception equally well.

correctly reflects this brute fact about where the hierarchy of sets happens to stop.

Moreover, the sets lose a substantial aspect of their appeal as a mathematical foundation if we can't capture all talk of coherent mathematical structures within set theory, i.e., all such structures can be realized in a set model[3]. However, it is (at best) unclear whether we can do this if we accept actualism and say that the hierarchy of sets doesn't 'go all the way up' in the sense indicated above. Of course, by Gödel's completeness theorem for first order logic, any consistent collection of first order axioms will have a model. However our conceptions of mathematical structures (like, famously, the natural numbers) can include non-first order notions, like second order quantification. So the completeness theorem doesn't guaranteed that our conceptions of these structures will have 'intended' models in the hierarchy of sets (i.e., models which treat their non-first order vocabulary standardly).

One might further press this objection, by arguing as follows. If there were an actualist hierarchy of sets we could refer to, then we could also uniquely describe the possible structure which you would get by adding a single layer of classes to this hierarchy of sets. This structure is a legitimate topic for mathematical investigation, and yet this structure is not instantiated anywhere within the hierarchy of sets.[4]

Now, we could avoid the above worry about arbitrariness while securing a definite height for the hierarchy of sets, by simply *adding* some new idea

[3]Note that such a model must actually realize the structure in question, i.e., be 'isomorphic', not merely realize the same first-order theory.

[4]See [44] Hellman for a version of this generality worry.

about height to our current conception of the hierarchy of sets. For example, we could say that the sets are the shortest possible structure satisfying ZFC_2 (i.e., the sets up to the first inaccessible). This proposal is somewhat analogous to saying that the numbers are 'as short as can be' while being closed under successor and satisfying all the other first order Peano Axioms – as we do when we take the natural numbers to satisfy induction. However, making this kind of height-minimizing stipulation seems to fit badly with actual mathematicians' interest in large cardinals (which require the set theoretic hierarchy to extend far beyond the shortest model of ZFC). And stipulating any height for the hierarchy of sets does nothing to help with the secondary worry above, that actualists shortchange the intended generality of set theory. If anything, it seems to make this problem worse.

2.3 Categoricity and Quasicategoricity Arguments

2.3.1 McGee and Appeal to Ur-elements

In 'How we learn mathematical language'[67] Vann McGee advocates an interesting and influential conception of an iterative hierarchy of sets with ur-elements which might seem to help solve the problem of commitment to an arbitrary stopping point for the hierarchy of sets noted above.

However I will argue that this is an illusion. Although McGee's characterization of a hierarchy of sets solves the problem he is concerned with in that paper (addressing a certain kind of referential skepticism), it does not make the height of the actualist hierarchy of sets look any less arbitrary.

In [68] Van McGee defends realist claims that we can secure definite reference to the hierarchy of sets up to isomorphism (and thereby justify our presumption that all questions in the language of set theory have definite right answers) from a reference skeptical challenge.

Specifically he proposes an account of how creatures like us could count as having a definite conception of the sets up to isomorphism, given the presumption that we can secure definite realist reference for other kinds of vocabulary, including (it will be important to note) presumption that we are somehow able to quantify over absolutely everything (sets included).

First McGee explains how we can secure (the effect of) definite reference to second order quantification and thus uniquely describe the intended width of the hierarchy of sets, via a story about schemas which I won't summarize here. Then he suggests that we can pin down the intended height of the hierarchy of sets by considering a conception of a hierarchy of sets *with ur-elements*.

The idea of set theory with ur-elements is simply that sets don't just have sets as elements. They can also have objects that aren't sets as elements. So there will be sets of elephants, electrons and spacetime points and the like. A common way of developing set theory with ur-elements is to keep the core idea of an iterative hierarchy of sets described above (with each layer containing 'all possible subsets' from the lower layers), but take the lowest level of the hierarchy of sets to include sets corresponding to all ways some way of choosing from among all the objects that aren't sets (e.g., elephants, billiard balls, electrons, marriages and the like), rather just the empty sets. That is (one might say) the hierarchy of sets with ur-elements starts from the plurality of objects that aren't sets, rather than starting from nothing. Note that the hierarchy of sets with ur-elements includes all pure sets. Thus, uniquely pinning down a hierarchy of sets with ur elements would suffice to pin down a hierarchy of pure sets as well.

McGee shows that we can (in a sense) pin down the intended height of this hierarchy of sets with ur-elements, if we accept the following axiom.

Urelement Set Axiom. $(\exists x)(Set(x) \wedge (\forall y)(\neg Set(y) \rightarrow y \in x))$

This axiom says that there's a set which contains, as elements, all the objects that aren't sets. And McGee proves that $ZFC_2 + U$ (the result of adding the above ur-element principle to second order ZFC set theory) has a property which he calls 'quasi-categoricity'[5]. Given any single choice of a total domain (what you are quantifying when you quantify over everything *including the sets*) there cannot be two non-isomorphic (with respect to $\in$) interpretations of set theory which both: choose 'sets' from within this domain, take quantifiers to range over this whole domain and make McGee's $ZFC_2 + U$ come out true (while interpreting all logical vocabulary standardly). So, for example, we couldn't have a single universe containing both a hierarchy of red sets and a hierarchy of blue sets, such that both hierarchies satisfy the constraints imposed by $ZFC_2 + U$ on their relationship to

[5] One might worry about the above axiom on the basis of Uzquiano's [105] proof that McGee's axioms for set theory with urelements are incompatible with certain axioms of mereiology, but I leave this question aside as the concerns I will be raising are unrelated.

the total universe (red sets and blue sets included).

In a nutshell, McGee's argument for quasi-categoricity goes like this. By standard results, any two set-sized models of pure second order ZFC of the same cardinality are isomorphic. And $ZFC_2 + U$ implies that the hierarchy of pure sets has the same cardinally as the total universe (sets included). For clearly it is possible to map the sets 1-1 into the universe by sending every set to itself. And McGee argues that it's possible to 1-1 map the universe into the collection of pure sets as follows. By the Ur-element Axiom, all the ur-elements form a set. And by the ZFC axioms every set can be well ordered, and the map g taking each ur-element to its ordinal height in this well-order is an injection from the non-sets to the ordinals. We can now use transfinite induction to define a 1-1 map H from the universe to the pure sets. $H(x) = \langle g(x), 0\rangle$ for x an Urelement. $H(x) = \langle \{H(y) : y \in x\}, 1\rangle$ for x a set.

Thus any two hierarchies both satisfying $ZFC_2 + U$ in the same universe (like the red sets and blue sets imagined in the example above) must have the same cardinality as the entire universe. So they must have the same cardinality as each other and hence be isomorphic.

I think this is a nice result, which does the job McGee wants done: answering skeptical challenges about definite reference to the hierarchy of sets, on behalf of a platonist who presumes that there is a definite heirarachy of sets which does stop somewhere and we can somehow unproblematically quantify over absolutely everything (these sets included). However establishing quasi-categoricity does nothing to address our current worry that actualists are committed to an additional and arbitrary joint in reality: a point where the hierarchy of sets just happens to stop.

McGee's theorem shows that if there is an actualist hierarchy of sets with ur-elements, and we use unrestricted quantification to refer to the hierarchy of sets uniquely (up to isomorphism) by saying that the sets have the same cardinally as the universe as a whole (i.e., the sets together with the non-sets).

But this fact that (assuming actualism and unproblematic quantification over everything) we are able to determinately refer to the height of the hierarchy of sets does not imply that we have any beliefs which logically necessitate (and thereby make non-arbitrary) facts about where the hierarchy of sets happens to stop. Indeed, as McGee himself points out, the conception of sets he articulates is **not** categorical; the beliefs about the sets which he invokes are compatible with many different possibilities about how large the total universe of sets is. It is only quasi categorical, in the sense that any two interpretations of the beliefs in this conception *which agree on the size of the total universe, sets included* (and all logical vocabulary) must interpret our set talk as referring to isomorphic structures.

One could use McGee's conception of sets with ur-elements in a slightly different way which *would* block the arbitrariness worries for actualism I've pressed above, as follows. Take our conceptions of all particular non-set objects (elephants, electrons, contracts etc.) to pin down the cardinality of objects that aren't sets. Then take the combination of that with our acceptance of McGee's ZFCU to pin down a hierarchy of sets formed from these ur elements. Then say that this hierarchy of sets stops *as soon as it can* while satisfying ZFC_2 and containing a set containing all these non-set objects. In doing this we would not have provided a categorical conception of the sets, but a quasi categorical conception which (unlike McGee's story) blocks the need for appeal to extra brute joints in reality regarding set theory. For we would have shown how the facts about what non-mathematical objects of various kinds exist combine with our conception of an iterative hierarchy of sets with ur-elements to pin down a unique structure (up to isomorphism). Unfortunately however, this proposal faces the same worries about making the hierarchy of sets too small, which arose for the idea that we could just pick a restrictive conception of the sets in section 2.2. It also makes the height of the hierarchy of sets contingent!

2.3.2 Martin

Similarly I should note that although Martin proves a kind of categoricity result for set theory in [66]. While this theorem may pose difficulties for anti-objectivist approaches, I want to take a moment to note that it doesn't block the indefinite extensibility/arbitrariness I've advocated above.

Indeed Martin seems to positively endorse worries about whether we have a definite conception of the height of the hierarchy of sets. For in [65] he distinguishes four ingredients in our conception of the hierarchy of sets as follows.

> The modern, iterative concept has four important components:

1. the concept of the natural numbers;
2. the concept of sets of x's;
3. the concept of transfinite iteration;
4. the concept of absolute infinity.

> Perhaps we should include a the concept of Extensionality as Component (0). Component (1) might be thought of as subsumed under the other three, but I will treat it separately.

And then he expresses the following reservations about whether we have a definite coherent concept of the fourth component 'absolute infinity' as follows.

> Cantor described the sequence of all the ordinal numbers as "absolutely infinite, so I am using the term "absolute infinity for the concept that is the fourth component of the concept of set. One can argue that the concept is categorical, and that any two instantiations of the concept of set (of the concept of an absolutely infinite iteration of the sets of x's operation) have to be isomorphic. But it is hard to see how there could be a full informal axiomatization of the concept of set. There are also worries about the coherence of the concept. People worry, e.g., that if the universe of sets can be regarded as a "completed totality, then the cumulative set hierarchy should go even further. Such worries are one of the reasons for the currently popular doubts that it is possible to quantify over absolutely everything. I am

also dubious about the notion of absolute infinity, but this does not make me question quantification over everything.

However, in [66] Martin proves there's a sense in which our conception of the hierarchy of sets is categorical, as part of an argument against plenitudinous anti-objectivist approaches to set theory (like Hamkins which we will discuss in [41]) in which we say that certain set theoretic claims Φ are not determinately true or false. The anti-objectivist justifies this belief by claiming there are many different (non-isomorphic) hierarchies of sets which all answer our conception of the total hierarchy of sets, and some of these make Φ true and others make Φ false.

Martin notes that if we accept a certain conception of the hierarchy of sets (and the principles below), it follows that there could not be two different non-isomorphic hierarchies of sets 'the red sets' and 'the blue sets' could not simultaneously exist.

- a 'uniqueness' principle: all sets are extensional. That is, if there are two distinct sets x and y (even in two different putative hierarchies!), then there must be some object which is an element of x but not y or vice versa. Thus, for example, there can be only one set $\{Mars, Venus\}$.

- a conception of the hierarchy of sets, including (among other more familiar elements) the following height closure principle: if a set exists, then any hierarchy of sets containing the elements of that set must contain the set itself

Martin shows that it follows from the principles above, essentially by induction, that there can't two different hierarchies of sets. Any two putative hierarchies of sets satisfying the conditions above, must agree on their urelements, and then on their first layer and their second layer etc.

One can call this a categoricity result. But note that it doesn't imply that it's, logically or metaphysically necessary that any collections of objects which satisfy our conception of sets must have a certain (unique) structure. Rather, it merely shows that there can't be two distinct *actual* set-theoretic hierarchies. For example, Martin's argument doesn't rule out the possibility that there could be some description of an ordinal ϕ_κ, such that it would be logically possible to have a structure satisfying our conception of the sets containing an ordinal satisfying ϕ_κ but also logically possible to have such a structure which didn't contain any ordinal satisfying ϕ_κ. Merely that we couldn't have two actual hierarchies of sets one of which contains ϕ_κ while the other does not.

2.4 A Problem Justifying Replacement

In addition to the worry above (about whether we have a coherent conception of the intended height of the hierarchy of sets), set theoretic Actualists also face a problem about justifying the axiom of replacement.

Informally, the axiom schema of replacement tells us that the image of any set under a definable (with parameters) function is also a set. More formally, let ϕ be any formula in the language of first-order set theory whose free variables are among $x, y, I, w_1, \ldots, w_n$. We can think of the formula $\phi(x, y)$ (and choice of parameters) as specifying a definable function (with parameters) taking x to the unique y such that $\phi(x, y)$. Then the instance of axiom schema of replacement for this formula ϕ says the following:

(Repl)
$$\begin{aligned} &\forall w_1 \forall w_2 \ldots \forall w_n \forall a \\ &\qquad \forall x(x \in a \rightarrow (\exists! y)\phi(x, y, w_1, ...w_n)) \\ &\qquad \rightarrow \exists b \forall x\big(x \in a \rightarrow \exists y(y \in b \wedge \phi(x, y, w_1, ...w_n))\big)\big] \end{aligned}$$

So replacement says: whenever some first order formula defines a function on a set A, i.e., associates each element x of A with a unique y, there is a set B equal to the image of A. In other words, the hierarchy of sets extends far enough up that all the elements in the image of A can be collected together.

As Boolos points out in [12], the axiom of replacement imposes a kind of closure condition on the height of the hierarchy of sets which doesn't obviously follow from the iterative hierarchy conception of the sets above. For consider $V_{\omega+\omega}$. This structure satisfies the iterative hierarchy of sets conception above; each layer of sets contains sets corresponding to all ways of choosing sets lower than it, and there isn't a last layer. However, it doesn't satisfy Replacement, since you could take the set ω (formed at layer $V_{\omega+1}$) and write down a function ϕ which associates 1 the $\omega + 1$, 2 with $\omega + 2$ etc. Then for each x in ω there's a y in $V_{\omega+\omega}$ satisfying $\phi(x, y)$. But there isn't any set b in $V_{\omega+\omega}$ which collects together the image of every member of ω.

That set b is only formed at a $V_{\omega+\omega+1}$. This raises a worry about how to justify Replacement, and (indeed) whether mathematicians are justified in using it at all.

So (even if we take for granted that there are objects satisfying the iterative hierarchy conception of sets), if we want to justify use of the ZFC axioms, a question remains about how to justify the axiom of replacement.

There has been much interest and sympathy with this worry in the subsequent literature. For example [82] Putnam writes, "Quite frankly, I see no intuitive basis at all for . . . the axiom of replacement. Better put, I do not see that a notion of set on which that axiom is clearly true has ever been explained." And in a discussion of the history of set theory Michael Potter remarks that, "it is striking, given how powerful an extension of the theory replacement represents, how thin the justifications for its introduction were" [6], and then reports that "In the case of replacement there is, it is true, no widespread concern that it might be, like Basic Law V, inconsistent, but it is not at all uncommon to find expressed, if not by mathematicians themselves then by mathematically trained philosophers, the view that, insofar as it can be regarded as an axiom of infinity, it does indeed, as von Neumann ... said, 'go a bit too far"[78].

[6]He supports this assessment by quoting "Skolem... gives as his reason that 'Zermelos axiom system is not sufficient to provide a complete foundation for the usual theory of sets, because the set $\{\omega, P(\omega), P(P(\omega)), ...\}$ cannot be proved to exist in that system; yet this is a good argument only if we have independent reason to think that this set does exist according to 'the usual theory, and Skolem gives no such reason. Von Neumanns ... justification for accepting replacement is only that, ' in view of the confusion surrounding the notion not too big as it is ordinarily used, on the one hand, and the extraordinary power of this axiom on the other, I believe that I was not too crassly arbitrary in introducing it, especially since it enlarges rather than restricts the domain of set theory and nevertheless can hardly become a source of antinomies."'.

To my knowledge, three main (actualist) strategies for justifying replacement are currently popular.

First, people have tried to justify using the axiom of replacement 'extrinsically' in much the way that one would justify a scientific hypothesis, by appeal to its fruitful consequences, arguing it helps prove many things we independently have reason to believe and hasn't yet been used to derive contradiction or consequences we think are wrong[7].

Second, people have tried to justify replacement by noting it follows from a set theoretic reflection principle, which says that for any finite list of formulas there is an initial segment of the hierarchy of sets, V_α, which behaves like the whole hierarchy of sets with regard to these formulas, i.e., for any objects $a_1, \ldots, a_n$ in V_α, we have $\phi(a_1, \ldots, a_n) \leftrightarrow V_\alpha \models \phi(a_1, \ldots, a_n)$. [8]

[7] See [52] which quotes Gödel's description, "Even disregarding the intrinsic necessity of some new axiom, and even in case it had no intrinsic necessity at all, a decision about its truth is possible also in another way, namely, inductively by studying its "success, that is, its fruitfulness in consequences and in particular in "verifiable consequences, i.e., consequences demonstrable without the new axiom, whose proofs by means of the new axiom, however, are considerably simpler and easier to discover, and make it possible to condense into one proof many different proofs."

[8] To see how this implies replacement, consider some instance of the replacement schema for some set a and formula ϕ. This is

$$[(\forall x \in a)\,(\exists! y)\phi(x, y)] \rightarrow [(\exists b)\,(\forall x \in a)\,(\exists y \in b)\phi(x, y)]$$

If the antecedent is false this is trivial. So suppose that $[(\forall x \in a)\,(\exists! y)\phi(x, y)]$. Now, let $\Phi(x, y)$ be the conjunction of the instance of the replacement scheme applied to ϕ and $\phi(x, y)$ and let β be large enough that for some $x, y \in V_\beta$ we have $\phi(x, y)$ and $a \in V_\beta$. By reflection we can infer there is some α such that

$$\beta \in \alpha \wedge (\forall x, y \in V_\alpha)\left[\Phi(x, y) \leftrightarrow \Phi^{V_\alpha}(x, y)\right]$$

Instantiating x, y with the x, y in $V_\beta \subset V_\alpha$ can infer that $\Phi^{V_\alpha}(x, y)$ holds and hence this instance of replacement holds in V_α. Let $b \in V_\alpha$ be the witness to this fact. Given x in a there is some y such that $\phi^{V_\alpha}(x, y)$ and by the instance of reflection above that entails $\phi(x, y)$ as desired.

Third, people have tried to justify replacement by appeal to a kind of inference to the best explanation along the following lines. Russell's paradox tells us that not all pluralities of objects can form a set (there isn't a set of all sets that aren't members of themselves). So if there are any sets, there should be a principled division between those pluralities of objects can form a sets and those which can't. But sets don't have that many features. So (one might think) size is the only natural choice for the limitation on what pluralities count as sets and it should be the *only* such limitation [78]. As Michael Potter puts it, we should accept the following Size Principle, "If there are just as many Fs as Gs, then the Fs form a collection if and only if the Gs do." (which implies Replacmenet) because

> "[A] collection is barely composed of its members: no further structure is imposed on them than they have already. So... what else could there be to determine whether some objects form a collection than how many there are of them? What else could even be relevant?"

However, even if these strategies succeed in providing *some* justification for using the axiom of replacement, I think it will be agreed that none of them account for the kind of intrinsic convincingness we usually expect (and hope for) from mathematical axioms.

The first strategy only provides extrinsic justification, whereas it's at least prima facie appealing to expect central principles of set theory which are used without comment to have intrinsic justification, and this expectation seems common in other areas of mathematics. For example, it seems that everything we want to say about the natural numbers follows from (say) our second order conception of the natural numbers. Now if it turns out that adequate intrinsic justification cannot be given, it might be reasonable to accept extrinsic justification (we do in the sciences after all). And probably we will reach a point with, e.g., large cardinal axioms where extrinsic justification is all we can provide, However, insofar as intrinsic justification can be provided for the ZFC axioms which are treated as quite secure and used to provide a foundation/explication of normal mathematical claims that we are very confident in, it would be desirable to provide an internal justification.

The second strategy (justification by appeal to a reflection principle) arguably is *somewhat* attractive. For, as Koellner reviews in one can motivate reflection principles[9] by Gödel's idea that the total hierarchy of sets (V) should be impossible to define. For reflection principles (in effect) say that anything that's true of the whole hierarchy of sets will also be true in some proper initial segment of it. If some instance reflection principle failed (so there was some fact about the whole hierarchy of sets that didn't reflect down to be true of a proper initial segments of the sets) then we could (in a sense) define the hierarchy of sets by saying it is the shortest[10] iterative hierarchy structure satisfying this claim. Gödel writes

> "Generally I believe that, in the last analysis, every axiom of infinity should be derivable from the (extremely plausible) principle that V is indefinable, where definability is to be taken in

[9]Different reflection principles correspond to different classes of sentences being reflected. For instance, you might think only first order sentences reflect or first order formulas with parameters or second order sentences etc.

[10]That is, the sets satisfy the non-reflected claim but no initial segment does.

[a] more and more generalized and idealized sense."[11]

However, although the idea in the quote above has a kind of elegance and intrinsic appeal, it's not obvious (or not as obvious as we'd naively hope foundational axioms for mathematics could be) that there could be a structure satisfying this principle together with our other expectations about the hierarchy of sets (e.g., the other ZFC axioms, and the width conditions discussed above). Also, to the extent that Gödel's idea in the quote above motivates the first order Reflection principle used to justify Replacement above, it would seem to also motivate third order reflection, some instances of which (as Koellner notes in the article cited above) have been shown to be inconsistent[90].

As regards the third strategy (Potter's inference to the best explanation), I will suggest that sets do have *some* other features than their size which could be used to give an explanation for why there isn't a Russell set analogous to the one Potter gives.

In particular note that, given the iterative hierarchy conception of sets (which Potter accepts) each set will have the property of first being generated at some ordinal level α. This feature of sets as a fairly natural and principled one. One can think of it as reflecting how many layers of indirect and metaphysically derivative object existence (given the common idea that sets are in some sense metaphysically dependent on their elements not vice versa [12]) one has to go through to arrive at that set.

[11] This is quoted from [106] in [52].

[12] See, for example [10] for the a development of the intuition that the existence of Socrates' singleton is to be grounded in the existence of Socrates and depends on that,

So, rather than hypothesizing (with Potter) that the iterative hierarchy of sets stops at a certain point because ascending any further would require collecting objects which are *too plentiful* to form a set, couldn't we just as well hypothesize that the iterative hierarchy of sets stops somewhere because any further sets form would have to occur *too high up* in an iterative hierarchy (i.e., one would have to ascend through too many layers of abstraction/metaphysical dependence to form a set from the relevant elements). To the same (rather fanciful) extent that we can imagine that the rubber band holding together the elements of a sets just happens to be too small to collect any plurality of elements of a certain size κ, we could imagine that the power of lower level sets to ground the existence of higher level sets and thereby indirectly to ground the existence of still higher level sets etc. eventually becomes too attenuated to allow any further sets to be formed at some height α.

One might also object to Potter's methodology more generally, on the grounds that even philosophers who are happy to use this kind of metaphysical inference to the best explanation suggested by potter's justification don't usually take applying this method to justify the great confidence and certainty we feel we give to typical mathematical results.

So how do we know there's an upper bound to the *sizes* sets can have vs. an upper bound to the *heights* they can have?

In order to justify the level of confidence we have in set theory, and par-

in a way that the existence of Socrates does not depend on the existence of his singleton, and use of this intuition to motivation a notion of grounding which is distinct from metaphysically necessary covariation and supervenience.

ticularly Replacement, (as well as for aesthetic reasons) we would like our set theoretic axioms to follow from some simple intuitive conception which strikes us as prima facie logically coherent. For instance, we think of number theory as describing the sequence built by starting at 0 and continuing to add successors 'as long as is needed to ensure that there is no last natural number, but no longer' in a sense which can be cashed out via the second order axiom of induction. And we can think of the reals as describing a line extending to infinity in both directions without gaps (i.e., such that it's impossible to add any further 'number' anywhere on the line without it being equal to a real[13]). In both these cases our conception of the mathematical structure seems to flow from a single unified conception that's intuitively consistent.

The iterative hierarchy idea offers such a conception for set theory without replacement (imagine an iterative hierarchy whose height stops at $V_{\omega+\omega}$), but we are left with only the relatively weak arguments above to convince us that the axiom of Replacement can be consistently and appropriately added to this conception.

I will suggest that moving to a potentialist approach to set theory lets us do better. It lets us provide clear and intuitively coherent notion of what set theory is about which suffices to justify all the ZFC axioms including Replacement.

[13]One can think of a Dedekind cut which doesn't correspond to a real number as a kind of gap, i.e., a vertical line passing through the x-axis that somehow misses every real number.

2.5 Indefinite Extensibility

I have suggested that we don't have a consistent categorical conception of the intended height of the hierarchy of sets. And in the chapters below I will develop a potentialist approach to the height of the hierarchy of sets.

However, it should be noted that other philosophers interested in potentialism have explored more general versions of potentialism, which go further and reject the idea that we have a definite conception of the structure of the natural numbers or the width of the hierarchy of sets. Thus one might wonder if there is a principled reason for taking a potentialistic approach to the height of the hierarchy of sets but not to the natural numbers or the width of the hierarchy of sets.

2.5.1 Height Potentialism And No More

In a nutshell, I think my stance is principled because our naive conception of the height of the hierarchy of sets gives rise to a Burali-Forti paradox while no similar problem seems to arise with our naive conception of the natural numbers, the width of the hierarchy of sets or full second order quantification.

More specifically, in the pages above, we've seen how one can fairly concretely imagine constructing an ordinal-like-object above any well ordered plurality of ordinals, and a layer of set- like-objects above any plurality of sets. We can specify exactly how $\leq$ and $\in$ would relate the new sets/ordinal to all the old sets/ordinals previously considered. And we have a kind of positive intuition that the new object we've imagined adding has a good a claim to be called a set/ordinal as the original plurality (whatever it may be). It's prima facie plausible that the structure we imagine forming by extending any given plurality of ordinals has as good a claim to contaning all the objects that satisfy our conception of 'the ordinals'/'the sets' the original structure. For our conception doesn't seem to include any (coherent) negative conditions which say that they height of the hierarchy must stop at a certain point.

But we can't do the same thing with our concepts of 'full' second order quantification (aka arbitrary subsets of a given collection), natural number and real number. Perhaps, in a sense it's intuitive that, for any collection of natural numbers (finite or infinite) we can imagine a strictly larger *vaguely* number-like objects .We can always imagine adding (something like) a successor or a limit ordinal after all numbers within any collection of numbers. However, our grasp of the natural numbers does very centrally include such a principle saying the numbers must stop at a certain point, namely the second order induction axiom! We think the numbers are (so to speak) as *few as can be*[14] while containing 0 and the successor of everything they include, and that for this reason any property which applies to 0 and applies to the successor of everything it applies to must apply to all the numbers. The same goes for the concept of full second order quantification/all possible subsets of a given collection. We have no positive intuition about how to

[14]Here I mean 'few' in an ordinal sense not a cardinal sense. Maybe it would be better to say that the natural number structure is as short/small as can be while satisfying this condition

generate, for any given collection of sets of cats a new set-of-cats like object which is distinct from all the ones previously consdered.

2.5.2 Contrast with Dummett

It may be helpful at this point to contrast my arbitrariness problem for actualism with Michael Dummett's influential arguments about indefinite extensibility in *The Seas of Language*. In that [25], Dummett raises something very much like the Burali-Forti worry I pressed above concerning the height of the hierarchy of sets. For example, when speaking about cardinalities and size rather than ordinalities and lengths of possible well orderings, he writes as follows.

> If it was... all right to ask, "How many numbers are there?", in the sense in which "number" meant 'finite cardinal', how can it be wrong to ask the same question when "number" means 'finite or transfinite cardinal'? A mere prohibition leaves the matter a mystery. It gives no help to say that there are some totalities so large that no number can be assigned to them. We can gain some grasp on the idea of a totality too big to be counted, even at the stage when we think that, if it cannot be counted, it does not have a number; but, once we have accepted that totalities too big to be counted may yet have numbers, the idea of one too big even to have a number conveys nothing at all. And merely to say, "If you persist in talking about the number of all cardinal numbers, you will run into contradiction", is to wield the big

stick, not to offer an explanation.[15]

And one might say that both of us reject standard actualist set theory on the grounds that our conception of sets is 'indefintie extensibility' . However Dummett is concerned with indefinite extensibility in a different sense than I am. Specifically, I reject standard (actualist) platonsim about set theory because our concept of sets and ordinals is 'indefinite extensibility' in the following strong sense:

> **Strong Indefinite Extensibility** We have a positive intuition that for any hierarchy of sets/ordinals *there could be* there could be a strictly larger one which matches our iterative hierarchy conception of sets/ordinals equally well.

In contrast Dummett rejects standard platonist set theory because our concept of sets is 'indefinite extensibility' in this weaker sense:

> **Weak Indefinite Extensibility** For any collection of them *we can definitely imagine* (which he says he will start by presuming means any *finite* collection) this collection can be extended so as to contain extra things which would also fall under our conception of that structure[16]

Thus Dummett's reason for worrying about the sets applies to the natural numbers and real numbers (any *finite* collection of these will be missing a

[15] [25] pg. 439

[16] Dummett writes, "[A]n 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, characterize a larger totality all of whose members fall under it."[25] pg 440].

number which could be added) etc. while (we've just seen above that) mine doesn't. Philosophically speaking, I take these different indefinite extensibility worries to arise from different philosophical projects and background assumptions as follows.

Prima facie I take both the naive intuition that we mean something definite by both 'all possible subsets' and 'all the way up' at face value, until Burelli-Forti paradox shows the latter is contradictory. Since no analogous paradox seems to arise for all possible subsets I'm happy to invoke this notion in expressing a conception of the natural numbers etc.

In contrast, Dummett starts for a more skeptical/cautious position and asks to be shown how one could 'convey' a definite concept of structures to someone who starts out only understanding finite collections. And he prima facie doubts that you could do so by, e.g., giving an operation like adding one and talking about closing under it or (as I would prefer) or appealing to a modal notion of 'all possible' subsets which applies to infinite collections[17].

[17]To support this reading note that Dummett argues that the concepts of natural number and real numbers are just indefinitely extensible by (seemingly) starting from the presumption that all totalities we can form a definite conception of are finite and the answers the concerns that it's question begging to do so as follows:

> Burden-of-proof controversies are always difficult to resolve; but, in this instance, it is surely clear that it is the other side that has begged the question. It is claiming to be able to convey a conception of the totality of real numbers, without circularity, to one who does not yet have it. We are assuming that the latter does not have, either, a conception of any other totality of the power of the continuum. He therefore does not assume as a principle that any totality of which it is possible to form a definite conception is at most denumerable: he merely has as yet no conception of any totality of higher cardinality. Likewise, a conception of the totality of the natural numbers is supposed to be conveyed to one as yet unaware of any but finite totalities; but all that he is given is a principle of extension for passing from any finite totality to a larger one.

Thus, I take it, the pages above express a principled reason for doubting that we have a consistent categorical conception of the hierarchy of sets which (unlike Dummett's concerns) doesn't similarly call into doubt the appearance that we have a definite categorical conception of the natural numbers, the real numbers, or of what it takes for some layer of classes to include classes witnessing 'all possible ways of choosing' from some definite collection[18].

[18]Now, (tangentially to the main argument above) Dummett would presumably challenge me to say how we could grasp the notion of 'all possible subsets' of a given infinite collection which we naively seem to grasp. And, in a nutshell, I'd answer that we can latch onto a notion of logical possibility which (we will see below) suffices to categorically describe the numbers and sets in the same way (whatever it is) that we can latch on to a notion of objective physical possibility/law. For example it might be that we get both notions by making certain core good inferences (e.g., the actual to possible 12.2.1 and uniform relabling 12.5.1 principles below the case of logical possibility, and some other kind of extrapolation in the case of physical possibility) which in a way under-determine which modal notion we mean and then benefiting from reference magnetism. Thus I suspect that Dummett's worry either (despite protests to the contrary) comes down to an argument from some principle of manifestability which would call reference to realist physical possibility/law facts into doubt as well reduces to mine or. However, I won't pursue this argument here because my present aim was only to explain how my worry differed from Dummett's not to answer his worry.

Chapter 3

Putnamian Potentialism: Putnam and Hellman

Let us now turn to potentialism, a different approach to set theory which promises to let us solve at least one of the two main issues for traditional actualist set theory above (the worry about arbitrary stopping points).

In a nutshell, potentialists interpret mathematicians who appear to be quantifying over the sets as really talking about the possibility and extendability of structures satisfying the iterative hierarchy conceptions of sets discussed above. We might say that potentialist translations talk about the possibility of there being (objects with the structure of) standard width initial segments V_α of the total hierarchy of sets V, and how some such initial segments could be extended by longer initial segments. They don't interpret set theorists as quantifying over any collection of existing objects, or even as talking about what follows from some axioms describing the supposed structure of the sets. Instead, they systematically interpret mathematical utterances which appear to quantify over the sets as having a much more complicated logical form.

Crudely speaking, the potentialist will interpret singly quantified existential claims $(\exists x)(\phi(x))$ in set theory (e.g. $(\exists x)(x = x)$), as saying (something like) that it's possible for there to be a standard width initial segment of the hierarchy of sets containing an object x satisfying ϕ (in this case $x = x$). And they will interpret set theoretic claims of the form $\forall x \phi(x)$ claims where ϕ is quantifier free, (e.g., $(\forall x)(\neg x \in x)$), as saying (something like) that it's necessary that any object x within a standard width initial segment of the hierarchy of sets has the property ϕ0.

What about set theoretic claims involving nested quantification? The potentialist will interpret statements of the form $(\forall x)(\exists y)\phi(x, y)$ (where ϕ has no quantifiers) as saying (something like) that it's necessary that for any standard width initial segment V and object x within it, it's possible to have larger initial segment V' containing and extending V, containing an object y. And the same pattern continues for more logically complex sentences.

But, unsurprisingly, there has been much philosophical disagreement and debate about how to fill the details of this picture out. For example, what is the correct notion of possibility to employ here? What does it take for some things to form a standard width initial segment? And what logical tools should we use to articulate our answer to the above?

In the next two chapters I will review the development of potentialist set theory by various philosophers within two major schools of potentialist set theory, noting some problems for each of these views which will motivate my own preferred version of potentialism.

3.1 Putnam

In [83] Hillary Putnam sketches a way of thinking about set theory in terms of modal logic: as talk about what 'models' of set theory are, in some sense, possible and how such models can be extended.

He introduces a notion of being a standard model of set theory, which is a model of set theory closed under subsets, i.e., a hierarchy of sets having standard width and no infinite descending chains under $\in$[1]. Putnam says that we can 'make this notion concrete' by thinking of models as physical graphs consisting of pencil points (or the analog of pencil points in space of some higher cardinality) and arrows connecting these pencil points. And he "ask[s] the reader to accept it on faith" that we can express the claim that some model is standard in this way "using no 'non-nominalistic' notions except the '$\Box$"' (where $\Box$ denotes the logical necessity operator).

With this notion of a concrete model in place, Putnam suggests that we can

[1]Specifically, Putnam writes "[A concrete] model will be called standard if (1) there are no infinite-descending 'arrow' paths; and (2) it is not possible to extend the model by adding more "sets" without adding to the number of "ranks" in the model. (A 'rank' consists of all the sets of a given-possibly transfinite-type. 'Ranks' are cumulative types; i.e., every set of a given rank is also a set of every higher rank. It is a theorem of set theory that every set belongs to some rank.)"

understand set theoretic statements as claims about what such models are possible, and how they can be expanded. For example, he proposes that we can paraphrase a set theoretic statement of the form '$(\forall x)(\exists y)(\forall z)\phi(x, y, z)$' where ϕ is quantifier free, as saying that, if G is a standard concrete model, and p is a point within G, then it is possible that there is a model G' which extends G, and a point y within G' such that necessarily, for any model G'' which extends G' and contains a point z, $\phi(x, y, z)$ holds within the concrete model G''. And we can treat arbitrary quantified statements in set theory in an analogous fashion.

Putnam then suggests that adopting this potentialist approach to set theory can help us dispel the kind of arbitrariness and indefinite extendability worries I discussed in section 2.2 above. For, adopting this approach lets us understand set theoretic talk without imposing or positing arbitrary limits on the size of structures (as we would do if we just stipulate a stopping point to the hierarchy of sets, or inferring that it must stop somewhere) in a way that seems faithful to our intuitions about the generality of set theoretic reasoning[2]. As Putnam puts it,

> "[W]e have a strong intuitive conviction that whenever As are possible, so is a structure that we might call 'the family of all sets of As.' ...from the standpoint of the modal-logic picture ... the Russell paradox ... shows that no concrete structure can be a standard model for the naive conception of the totality of all sets; for any concrete structure has a possible extension that

[2]In particular, (before thinking about the paradoxes) we'd hope for set theory to be general in the sense that every possible structure will have a copy somewhere in the sets.

> contains more "sets." (If we identify sets with the points that represent them in the various possible concrete structures, we might say: it is not possible for all possible sets to exist in any one world!) Yet set theory does not become impossible. Rather, set theory becomes the study of what must hold in, e.g., any standard model for Zermelo set theory."

Puntam's sketch offers an appealing style of response to the worries about arbitrary stopping points for the hierarchy of sets indicated above – one that (we will see) has inspired many other philosophers. However, this proposal is (explicitly) sketchy on certain formal and philosophical points. For instance, Putnam doesn't provide any criteria for what it would take for some collection of arrows and pencil points to form a standard model and he asks the reader to "accept it on faith that the statement that a certain graph G is a standard model for Zermelo set theory can be expressed using no 'non-nominalistic' notions except the '$\Box$'."[3]

And, philosophically, Putnam says very little about the notion of 'mathematical possibility' which he intends to capture with the $\Box$ and seems to vacillate between a purely mathematical understanding of necessity and a physical understanding.

For example he writes (brackets in original), "assuming that the notions of mathematical possibility and necessity are clear [and there is no paradox associated with the notion of necessity as long as we take the '$\Box$'s a statement

[3]Here nominalistic notions are ones that aren't committed to the literal existence of mathematical objects.

connective (in the degenerate sense of "unary connective") and not...as a predicate of sentences], I wish to employ these notions to try to give a clear sense to talk about 'all sets." However, at earlier points Putnam talks like 'conceivable' constraints on how many physical objects like pencil points and lines could fit into physical stance are relevant, and makes assumptions about this which philosophers like Parsons[74] and Tait[102] have been unwilling to grant, "I assume that there is nothing inconceivable about the idea of a physical space of arbitrarily high cardinality; so models of this kind need not necessarily be denumerable, and may even be standard."

Additionally, Putnam advocates potentialism as merely one possible and helpful 'perspective' on mathematics and claims that it is in some sense equivalent to a more familiar actualist understanding of set theory, which only appears to be incompatible with it. And cashing this idea out clearly requires serious and disputable metaphysics[4].

Furthermore, it's not clear that saying both perspectives are equally good is compatible with honoring Putnam's potentialism-motivating intuition that "whenever As are possible, so is a structure that we might call 'the family of all sets of As". We seem forced to *either* say that the idea that for any structure there could be a larger one is only true 'from the potentialist perspective' on mathematics or to say that it is true simpliciter, even from the actualist perspective. The former position can feel a little mysterious and unsatisfying while the latter is uncomfortable for two reasons. First (like more straightforward forms of actualism) it involves positing arbitrariness in

[4]See, for example, John Burgess' vigorous objections to Putnam's stance in [13].

mathematical reality by saying the actualist hierarchy of sets just happens to stop somewhere, though it could go on further. Second, its not clear (even at a very loose intuitive level) how talking about any such structure could be equivalent to a practice of modal set theory which considers arbitrary logically possible extendability[5].

3.2 Hellman

In [44] and [38] Hellman develops Putnam's ideas about potentialist set theory as part of a larger purely nominalist philosophy of mathematics – dropping Putnam's suggestion actualist and potentialist approaches to set theory are (somehow) supposed to be two equally good perspectives on the same thing.

Hellman dispels some of the unclarity about Putnam's idea of mathematical possibility by invoking a primitive notion of logical possibility. But he does relatively little to describe this notion. He does say that, "[when evaluating logical possibility] we are not automatically constrained to hold material or natural laws fixed." So it may be logically possible that $(\exists x)(\text{pig}(x) \wedge \text{flies}(x))$, but physically impossible. And he adds that, "we are free to entertain the possibility of additional objects even material objects of a given type.", which allows us to say that it's logically possible for there to be infinitely many objects even if there are only finitely many objects. This

[5]Perhaps one could say that the actualist hierarchy is the smallest standard width structure whose truth conditions for all first order logical claims agree with those provided by the potentialist set theory.

(arguably) lets us avoid concerns about limitations on the cardinality of space unduly limiting the range of possible models considered above. Beyond this remark, however, Hellman just suggests that his applications of logical possibility will make the notion he has in mind clear.

Hellman also does a lot to fill in the other promissory notes left by Putnam's sketch. He cashes out Putnam's appeal to 'standard models' of set theory by saying that standard models are models which satisfy ZFC_2 (i.e., the version of standard ZFC set theory which replaces the inference schemas of replacement and comprehension with corresponding second order axioms)[6]

Hellman replaces Putnam's claims about the possibility of concrete models where first order relations apply so as to satisfy certain descriptions with claims about the possibility of second order collections.

So, in particular, Hellman takes set theorists' singly-quantified existence claims, like $(\exists x)(x = x)$, to really be saying that it would be possible for a collection of objects V_0 to satisfy (a version of) ZFC_2 while containing a suitable object x (in this case, an x such that $x = x$).

More specifically[7], Hellman defines the claim that some second order collection X and relation f, (X, f) form a natural model of set theory as follows.

[6] So, for example ZFC contains the inference schema of comprehension, which has an instance for each formula in the language of set theory ϕ, saying that for every set x and choice of parameters $w_1...w_n$ there is a set $y = \{z \mid z \in x \wedge \phi(z)\}$ i.e. y such that $(\forall z)(z \in y \leftrightarrow z \in x \wedge \phi(w_1, ...w_n, z))$.

In contrast, by using second order logic one can state a single comprehension axiom as follows $(\forall x)(\forall C)(\exists y)(\forall z)(z \in y \leftrightarrow z \in x \wedge C(z))$. The same goes for the first order axiom schema of replacement and its second order analog.

[7] I suppress one detail of Hellman's paraphrase strategy (his separate treatment of set theoretic statements involving only restricted quantification) which makes no difference to the philosophical arguments being made here. See [44] chapter 2 section 2.

$$(X, f) \text{ form a natural model of set theory iff } ZFC_2^X\binom{\epsilon}{f}$$

where the expression on the right hand side of the biconditional is what you get by starting with ZFC_2 and then uniformly replacing all occurrences of ϵ with f, and reinterpreting all quantifiers as ranging over the objects in X rather than the sets. He then paraphrases singly quantified set theoretic statements (i.e. those of the form $\exists x\phi(x)$ were ϕ is quantifier free) as

$$\Diamond(\exists X)(\exists f)[(X, f) \text{ form a natural model of set theory } \rightarrow (\exists x)\phi(x)^X\binom{\epsilon}{f})]$$

For readability he then uses quantification over variables of the form V_i to abbreviate quantification over X_i, f_i which form a natural model of set theory, with claims of the form $z \in V_i$ standing for the claim that $z \in X_i$, for the relevant X_i. So, for example, the paraphrase for $\exists x(x = x)$ gets written as

$$\Diamond(\exists V)(\exists x)(x \in V \wedge x = x)$$

Similarly, he takes set theorists' singly quantified universal statements like $(\forall x)(x = x)$ to really say that it is necessary if X, f pick out a collection of objects satisfying ZFC_2 then any x chosen from among these objects would satisfy the version of ϕ which talks about this X, f structure (i.e. $\phi(x)^X\binom{\epsilon}{f})$. So $(\forall x)(x = x)$ will gets translated as

$$\Box(\forall V)(\forall x)(x \in V \rightarrow x = x)$$

Hellman handles nested quantification using claims about how collections of objects satisfying a version of ZFC_2 could be extended. Specifically, let us say that a model of set theory $V_2 = (X_2, f_2)$ extends another model $V_1 = (X_1, f_1)$ (written $V_2 \geq V_1$) iff X_1 is a subclass of X_2 and f_1 is the restriction of f_2 to X_1. This amounts to requiring that (X_1, f_1) and (X_1, f_2) pick out structures which relate to each other like initial segments of an actualist hierarchy V_α and V_β where $\alpha \leq \beta$.

Then Hellman would potentialisticly paraphrase the set theoretic sentence $(\forall x)(\exists y)(x \in y)$ as saying the following. Necessarily if V_1 satisfies ZFC_2 and includes a set x, it is logically possible for there to be an extension V_2 of V_1, also satisfying ZFC_2 and containing a set y such that $x \in y$ (in the sense of $\in$ relevant to V_2). Writing this out formally and using $\geq$ to say that one model of ZFC_2 extends another, we get:

$\Box(\forall V_1)(\forall x)[x \in V_1 \rightarrow \Diamond(\exists V_2)(\exists y)(y \in V_2 \wedge V_2 \geq V_1, \wedge x \in y)]$

3.2.1 Difficulties With Hellman

However Hellman's story has raised a number of objections.

The first of these (pressed by Hellman himself in [38]) concerns second order logic and Quine's complaint that second order logic is 'just set theory in disguise' hence ontologically committal in a way is incompatible with nominalism. When Putnam talks about the modal perspective on mathematics, he considers possibility and necessity claims about states of affairs involving specific first order relation vocabulary (be it 'number' or 'cat') [8]. However, as noted above, Hellman interprets set theoretic claims purely in terms of second order quantification. That is, instead of saying something about possible scenarios in which 'penciled point' and 'arrow' to apply as if to a standard model, we talk about how it is possible for some second order objects X and f apply as if to a standard model.

This creates a worry about ontological commitments because the second order paraphrases above will only capture (prima facie) intended/intuitive truth conditions claims about set theory if we suppose that a second order comprehension principle applies (with logical/metaphysical necessity). For example, we need to accept the idea that (as a matter of logical/metaphysical necessity) whenever there are first order objects satisfying some non-modal property, there is a second order object which collects together exactly these

[8] For example on pages 10-11 of [80] he writes" Let 'AX' abbreviate the conjunction of the axioms of the finitely axiomatizable subtheory of first-order arithmetic just alluded to. Then Fermat's last theorem is false just in case '$AX \supset \neg$ Fermat' is valid, i.e., just in case

(I) $\Box(AX \supset \neg$ Fermat)

Since the truth of (I), in case (1) *is* true, does not depend upon the meaning of the arithmetical primitives, let us suppose these to be replaced by "dummy letters" (predicate letters). To fix our ideas, imagine that the primitives in terms of which AX and $\neg$Fermat are written are the two three-term relations "x is the sum of y and z" and "x is the product of y and z" (exponentiation is known to be first-order-definable from these, and so, of course, are zero and successor). Let $AX(S,T)$ and $\neg$Fermat(S,T) be like AX and $\neg$Fermat except for containing the "dummy" triadic predicate letters S, T , where AX and $\neg$Fermat contain the constant predicates "x is the sum of y and z" and "x is the product of y and z." Then (1) is essentially a truth of pure modal logic (if it is true), since the constant predicates occur "inessentially"; and this can be brought out by replacing (1) by the abstract schema: (2) $\Box[AX(S,T) \supset \neg\text{FERMAT}(S,T)]$ -and this is a schema of pure first-order modal logic."

objects. Thus (if we take second order quantifiers to be ontologically committal, as Quine as well as Hellman circa [38] are inclined to do) we need to accept the existence of objects which the second order objects and function quantifiers $(\exists X), (\exists f)$ range over[9]. Thus, adopting Hellman's version of potentialist set theory can seem to impose significant ontological commitments[10].

Additionally, there's a worry about justification. Hellman demonstrates that his system vindicates standard first order (and bounded second order) reasoning about the sets. However, the principles he invokes to do this are hardly obvious. For instance the strong axiom of extendability which Hellman uses just brutely assumes the translation of replacement into a potentialist context and, indeed, Hellman doesn't claim anything like intuitive obviousness for this principle.

Hellman does provide a kind of 'external' justification for the use of the ZFC axioms on his version of potentialism. Hellman's justification goes like this. Assume that actualist set theory is true and there are cofinally many inaccessible cardinals. On this assumption, we can re-interpret (Hellman's preferred version of) potentialist claims as claims about what initial segments of the true hierarchy of sets exist. Then it is a theorem that, for each

[9] For example the second order comprehension principle needed to do potentialist set theory will also commit you to inferring from the existence of some mushrooms to the existence of a set of mushrooms (or other second order object with mushrooms as elements).

[10] In a paper published after the book [38], Hellman considers using mereology and plural quantification to replace second order quantification in reconstructing some of mathematics. However he doesn't attempt to extend this story to set theory. And even his reconstruction of analysis depends on simultaneously appealing to a logical (rather than metaphysical) possibility to secure the possibility of sufficiently collections of objects and then assuming that the laws of mereology (which are generally taken to be metaphysically rather than logically necessary laws) apply within all logically possible contexts.

first order set theory sentence ϕ, this re-interpretation of the potentialist translation of ϕ will be true iff the original sentence ϕ is true. Thus, since ZFC_2 is presumably true of the actualist hierarchy of sets, the potentialist translation of these claims will also come out true.

However (as Hellman himself notes), the justification he provides for the ZFC axioms is not satisfactory from a potentialist point of view, because it requires that we assume the existence of an actualist hierarchy of sets. Additionally, we must also assume that this hierarchy satisfies a further (somewhat) controversial large cardinal axiom: that there are co-finally many inaccessible cardinals (so, even from an actualist point of view, you might say that it justifies the more obvious on the basis of the less obvious). So, while Hellman's justification might be a useful rhetorical tool for convincing actualists, it doesn't provide a justification for using the ZFC axioms which the potentialist can accept.

Chapter 4

Previous Approaches: Parsons and Linnebo

4.1 Linnebo

So much for Putnam's potentialist set theory and Hellman's development of it. I will now discuss a philosophically different approach to potentialist set theory developed by Charles Parsons[71][73][74] and then Øystein Linnebo[59][60] [61] and lately Roberts [94][93].

We saw that Putnam and Hellman's formulations of potentialist set theory above appeal to a general notion of logical or logico-mathematical possibility which is supposed to constrain the behavior of all objects. And, indeed, the potentialist paraphrases they propose wind up completely eliminating apparent use of concepts like 'set' and 'element' in favor of claims about the logicomathematical possibility of non-mathematical first order relations like ‘point’ and ‘there is an arrow from... to...’ applying a certain way or there being second order collections and relations satisfying certain conditions.

In contrast, Linnebo (and Parsons and Roberts) take ‘set‘ and ‘element’ to be meaningful notions, and state various metaphysical principles about the nature and essence of sets (just as much as the platonist does). They then appeal to subject matter specific facts about the nature of sets in understanding and evaluating set theoretic claims. Additionally Linnebo invokes a special (apparently set theory specific) notion of what sets ‘could be generated’ or (more recently) ‘interpretational possibility’, rather than the general notions of logical or logico-mathematical possibility invoked by Putnam and Hellman.

Linnebo characterizes how his preferred ‘Parsonian’ approach to potentialist set theory differs from Putnam and Hellman’s potentialism as follows.

> “[On a Parsionian approach to set theory] the idea is not to ‘trade in one’s mathematical objects in favor of modal claims about possible realizations of structures but rather to locate some modally characterized features in the mathematical objects themselves. The mathematical universe is not ‘flat’. Rather, some of its objects stand in relations of ontological dependence, and the existence of some of its objects is merely potential relative to that of others.
>
> ‘A multiplicity of objects that exist together can constitute a

> set, but it is not necessary that they do. Given the elements of a set, it is not necessary that the set exists together with them. [. . .] However, the converse does hold and is expressed by the principle that the existence of a set implies that of all its elements.' (Parsons, 1977, pp. 2934)
>
> As Parsons emphasizes, this approach can also be used to explicate the influential iterative conception of sets, which tends to be explained by suggestive but loose talk about a 'process' of 'set formation'. It would be better, Parsons claims, to replace this talk of time and construction with 'the more bloodless language of potentiality and actuality."

Linnebo then formulates potentialist paraphrases for set theory which looks rather similar to Hellman's paraphrases. However, where (as we saw above) Hellman would potentialistically paraphrase the set theoretic sentence "$\forall x \exists y x \in y$", as follows:

$$\Box(\forall V_1)(\forall x)[x \in V_1 \rightarrow \Diamond(\exists V_2)(\exists y)(y \in V_2 \wedge V_2 \geq V_1, \wedge x \in y)]$$

Linnebo would potentialistically paraphrase this sentence as something more like[1]:

$$\Box(\forall x)[set(x) \rightarrow \Diamond(\exists set(y) \wedge x \in y)]$$

In both cases $\exists x$ claims are replaced by $\Diamond\exists x$ claims and $\forall x$ claims are replaced by $\Box\forall x$. However, there are three important differences.

[1] See page 213 of [60], but I have tried to make implicit quantifier restriction to sets in the language of set theory explicit.

First, note that Hellman builds a description of (standard width initial segments of) the iterative hierarchy of sets into his paraphrase (via his description of what it takes to be a V_i). In contrast, Linnebo takes (something like) the iterative hierarchy conception of sets to follow from necessary truths about pluralities and sets, like:

- plural comprehension:$\exists xx \forall u[u \prec xx \leftrightarrow \phi(u)]$,
- whenever there's a plurality xx of objects, there could be a corresponding set (i.e., a set whose elements are exactly the members of the plurality)
- sets and pluralities have their elements necessarily

Second, Linnebo uses plural quantification rather than second order quantification to express the idea that there could be sets corresponding to 'all possible ways of choosing' from ur-elements or sets below a certain level. He allows impredicative comprehension over all the pluralities existing in a certain scenario[2]. And he takes pluralities to have their elements necessarily.

Third (as noted above), Linnebo uses the $\Box$ and $\Diamond$ to express a 'subject matter specific' modal notion of what sets could be constructed (or what is 'intepretationally possible ' with regard to set theory). This notion of constructability is supposed to satisfy the following principles: whenever a

[2]See pages 210-211 of [62] "We adopt the following plural comprehension scheme: $\exists xx \forall u[u \prec xx \leftrightarrow \phi(u)]$, (P-Comp)...It should also be noted that, since $\phi(u)$ may contain bound plural variables, our plural comprehension scheme is impredicative. Impredicative plural comprehension can be motivated and justified by what, following Bernays (1935), we may call a 'quasi-combinatorial conception of pluralities. The idea is to extrapolate from the finite to the infinite. Just as we can run through a finite plurality, making arbitrary choices as to which elements are to be included in a subplurality and which are not, we can idealize and assume this to be possible for any plurality."

plurality of objects exists, a set of precisely those objects can be constructed. So, for example, the total collection of sets which has been constructed 'at a given point' must always be 'closed down' in the sense that if a set exists then everything that is (or ever could be) an element of that set also exists. Accordingly, Linnebo understands set theory as ultimately an investigation of what pluralities xx of (possible) objects could exist together in a single universe/scenario, and what sets could be 'formed' from these pluralities (with each set y necessarily having as members exactly those objects x belonging to this plurality xx as members, and existing only when these objects exist).

But how exactly shall we understand Linnebo's core modal notion (the 'could' invoked above)? In the next section I'll discuss two different ways of understanding the $\Box$ and $\Diamond$ in Linnebo's potentialist set theory, and explain my concerns about each. I will begin with a notion of possible social construction, and then consider an appeal to possible interpretations which Linnebo actually favors.

4.2 Cashing out Parsonian Set Theoretic Modality

4.2.1 The Constructivist Option

In papers like [59] and [60], Linnebo associates his modal notion with talk of constructability. And he invokes an image of a hierarchy of sets being constructed in time.

If set theory is really about what sets could be formed, is there a fact about what sets have *actually* been formed? It's hard to say no, because if we take the predicate 'set' to have a definite meaning[3], it seems that we should be able to ask what sets actually exist. The alternative of saying sets are special weird kind of object where it's meaningless to ask how many there are but only about how they could be - seems unattractively revisionary or at best very under explained. But if we say 'yes this is a meaningful question', then it seems that we need a story about what set construction involves, which could motivate a principled answer to the question of how high the hierarchy of sets currently goes up. Otherwise we seem committed to positing an arbitrary and unmotivated joint in nature (how many sets there happen to be).

One possible answer to the question of how sets are constructed, would be to say that you construct new layers of the hierarchy of sets by clearly imagining them. But then, as Parsons notes at the beginning of [75], it seems like even an infinite lifespan might not allow one to imagine oneself even moderately far up layers of a hierarchy of sets. Also one can't visually imagine adding a layer of sets above V_ω in the way that (maybe) one can always imagine adding an extra stroke to a series of visually imagined tally lines.

One could solve this problem by instead saying that we construct sets by doing something like accepting (first order?) set theoretic principles which require the existence of at least that many sets. So perhaps the mathemat-

[3] See the next subsection for an interpretationalist construal of Linnebo's set theory which might be seen as denying this.

ical community's (or individual mathematicians') 20th century decision to add axioms of infinity and replacement each amounted to constructing new objects extending the hierarchy of sets, because they increase the size of the smallest (well founded, since by Linnebo's principles the hierarchy of sets must all be well founded) hierarchy of sets that satisfies these principles.

But arguably this picture requires giving a kind of uncomfortable double interpretation of set theoretic talk. For consider someone who has accepted the ZFC axioms, and thereby socially constructed sets up to (i.e., all the sets below) the first inaccessible cardinal. Now suppose that they entertain (and wonder about but don't yet accept) a proposition whose (actualist truth) would require that there be more sets than have hitherto been constructed, e.g., the claim that there is an inaccessible cardinal.

On the story just sketched above, there are two different ways of associating truth conditions with this sentence. If one takes their word 'set' to refer to sets, and 'element' to relate objects to the sets they are elements of, then this claim comes out false. For no set with the relevant properties has yet been constructed (or, if we suppose that all consciously life is snuffed out a moment later, ever will be constructed). On the other hand, if one interprets this set talk potentialistically, as a claim about what sets could be constructed, this sentence will be true.

Thus mathematicians' talk of "sets" is given a kind of strange double duty. On the one hand, certain objects (actualist-sets) are brought into existence by mathematicians' talk, and count as "sets". But mathematicians' talk about "sets" never refers to these objects, but only abbreviates modal claims about which of them could be brought into existence.

To highlight what can seem strange about this in a different way, note how different this is from what we normally say about socially constructed objects. Maybe it's natural and normal to say that judges' talk of companies and marriages (and enforcing certain behavior regarding them) somehow brings these objects into existence by social constructions. But in these cases we expect that the truth of apparent existence claims about marriages and companies etc. to require that these objects exist (i.e. that they have actually been constructed)[4]– not merely that it would be possible for some such object to be constructed.

Another problems for this view concerns the relationship between socially constructed sets and facts about time and metaphysical possibility.

Should we say that sets are literally brought into being when the first person constructs them, so that (if the above idea that one constructs sets by accepting set theoretic axioms that require their existence is true) there were no sets before set theory was invented, and only sets up to V_ω at some point in the late 19th century and then more when Woodin came on the scene. The radical idea of a hierarchy of sets which comes into existence with human mathematical hypotheses and grows as postulates are made (so that presumably there would have been no sets if there had been no people) seems to be at least entertained by Linnebo and Hamkins at the beginning of a joint paper [42]. For example they describe potentialism in temporal

[4]Perhaps with some kind of quantifier restriction to those the further property of being recognized by the relevant court, since one country need not recognize another country's companies and marriages

terms as follows (italics mine).

> Set-theoretic potentialism is the view in the philosophy of mathematics that the universe of set theory is never fully completed, but rather unfolds gradually as parts of it increasingly come into existence or become accessible to us. On this view, the outer or upper reaches of the set-theoretic universe have merely potential rather than actual existence, in the sense that one can always form additional sets from that realm, as many as desired, but the task is never completed. For example, height potentialism is the view that the universe is never fully completed with respect to height: *new ordinals come into existence as the known part of the universe grows ever taller.*

And perhaps one could supplement this view with an error theory about why it seems strange to say that sets only came into being in the 19th century, as follows. Maybe it sound odd to deny that sets exist necessarily and timelessly because (as noted above) 99% of apparent claims about set existence should really be understood as expressing some potentialist paraphrase, and the potentialist paraphrase of the claim that some sets exist really is a timeless necessary truth.

Alternately, one might reconcile the idea of layers of the set theoretic hierarchy being constructed by mathematicians' acts in time with common opinions that mathematical objects are necessary and timeless, by drawing on some ideas from Cole[20] and Searle[95] about social construction and the possibility of decisions (about when a company came to exist, or when a player first qualified as on the injured list) taking effect retroactively[5].

4.2.2 The Interpretational Option

Linnebo advocates and develops a different, preferred, way of understadning the Parsonian potentialist $\Box$ and $\Diamond$ in papers like [61]. Here Linnebo invokes a modal notion which he calls interpretational possibility, which he distinguishes from both metaphysical, logical and mathematical possibility.

More specifically, in [61] Linnebo reviews (and seems to endorse) Parsons' arguments that we can't use any of the four different kinds of non-epistemic modality below to articulate potentialist set theory.

1. Physical or natural
2. Metaphysical or broadly logical
3. Mathematical modality
4. Logical modality in the strict sense.

Parsons says (and Linnebo agrees) that one can't use the notion of phys-

[5] Both have noted that public bodies like human rights courts or sporting associations seem to be able to assign social statuses like rights and being on the injured list retroactively, and similarly that they seem to be able to bring things like corporations into existence retroactively. Both have used this to suggest that objects which are -in a sense- contingently socially constructed at a certain time (e.g., the time of the court decision) might nonetheless be neccessarilly and eternally true.

But I think various philosophical work would need to be done to suitably develop this idea. For example: how are we to understand the relevant dependence of a necessary object on a contingent court decision, given that we cannot cash it out by saying that if the court had not made that decision then the right would not exist? It can also seem metaphysically odd and revisionary to suppose that the existence of a right or corporation in one possible world could be (ultimately and unavoidably) partly grounded in the facts about the verdicts of a court at (the actual one).

ical or metaphysical possibility to cash out potentalist set theory because all sets exist necessarily in this sense. And Parsons understands the notion of mathematical modality in terms of dropping 'all constraints of a metaphysical nature' and considering only what is 'compatible with the laws of mathematics' (in a way that is taken to include the existence of however many sets there are). So we also can't cash out a non-trivial potentialist sense in which it's contingent how many sets there are in terms of mathematical modality. And "the final notion of 'logical modality in the strict sense.. is fairly quickly set aside by Parsons, who finds it to be 'either … an awkward notion generally or not in the end [different from] from mathematical modality"'

Instead, Linnebo proposes that we should appeal to a notion of "interpretational possibility", which is distinct from all the notions above when formulating modal paraphrases. I take the thought to be something like the following. Even though whatever sets exists do so with metaphysical necessity, there are different ways of interpreting our talk of sets. However, I will now suggest that this key notion of interpretational possibility can't work like normal appeals to acceptable interpretations in various important ways.

For example, Linnebo can't mean interpretional possibility in the familiar Tarskian sense where all interpretations choose their domains from among some fixed universe of objects, otherwise we will have a maximum size which all interpretations of the sets have to be found within. On such a view actualists apparent commitment to an arbitrary stopping points, which potentialism promised to let us avoid, seems to get dragged back in.

Perhaps one could charitably interpret Linnebo's interpretational possibility in a quantifier variantist sense, where you get to choose between arbitrary logically possible structures for candidate disambiguations of what sets there are (rather than just choosing which objects from a fixed maximal domain are to count as sets). And indeed his independent interest in quantifier variance makes this possibility especially interesting. But, in this case, it seems to me that Linnebo's understanding of the interpretational possibilities relevant to set theory would have to draw on a prior understanding of something very much like the (general, subject matter neutral) notion of logical possibility which Putnam and Hellman invoke, operating in the background. For here we have an idea of general logico-combinatorial constraints (something like Hellman's notion of logical possibility) governing how a charitable interpreter could interpret someone's language by picking a possible precise quantifier sense which (in effect) associates each metaphysically possible world into a possible/logically coherent structure of objects which would count as existing in the relevant sense at this world.

Also, normally when there are many acceptable interpretations of how, say, the word 'martini' applies we say that it is ambiguous or vague whether a given borderline case is a martini. But, I take it, Linnebo doesn't want to say that it's *ambiguous or vague* whether, e.g., there's an infinite set. Instead he wants to say that this claim (i.e., the proposition expressed by saying that 'There's an infinite set' in mathematical contexts) is simply true, because its potentialist formalization is true.

We can also note that because Linnebo's key modal notion is explicitly stated to satisfy the converse Barcan-Marcus formula[6] and therefore excludes the possibility of worlds containing fewer objects than the actual world (indeed, the possibility operator only allows the space of objects to grow). Therefore, even in a purely formal sense, this can't be the independently motivated notion of logical possibility/validity discussed in Field [33] and section 6 below. Instead, he appeals to a sui generis notion of constructibility (or later interpretational modality), which exists alongside metaphysical possibility.

Additionally, Linnebo faces a version of the general justification problem raised for Putnam above. He does write down a set of axioms about which suffice to justify use of the ZFC axioms [60]. Thus (as I also will attempt to do) he provides a foundation for justification for using the ZFC axioms of set theory, which a potentialist can accept.

However, like, Hellman, Linnebo doesn't claim any strong intuitive obviousness for all the axioms he uses in this justification. When it comes to the modal principle which justifies use of replacement in his system, he brutely adopts an axiom that makes some pluralities capacity to form a set a matter of size. He doesn't provide further justification for this principle, or seem to think it is intrinsically obvious/deeply attractive. Instead he merely notes that similar assumptions have been proposed elsewhere in the literature on indefinite extensibility.

I take securing such a clear and intuitive interpretation of the potentialist's $\Box$ and $\Diamond$ to be a crucial tool for my current project of (potentialistically) jus-

[6] That is

$$\Box \forall x \phi(x) \rightarrow \forall x \Box \phi(x)$$

tifying core set theoretic principles (including powerful ones like replacement whose justification has seemed to pose philosophical problems for actualists).

For unless we have a crisp and intuitive understanding of what the potentialists $\Box$ and $\Diamond$ are supposed to mean, we are unlikely to have strong intuitions that further powerful axioms and inference rules (beyond what Linnebo has explicitly stipulated) involving this modal notion are correct. One might try simply *stipulating* that strong principles needed to justify potentialist replacement are supposed to be true as a constraint on /and explanation of what we mean by ‘constructibility’/‘interpretational possibility’. But (given our lack of any clear intended interpretation which makes these formal principles true) there will be little reason for thinking that the relevant principles are even syntactically consistent much less that they succeed in defining a suitable modal notion.

Instead, it would be better to lock on to an antecedently clear and meaningful modal notion, such that we have strong intuitions about how this notion applies, and then show that potentialist translations of all the ZFC axioms can be derived from intuitively true principles involving this modal notion. Doing this we would also justify confidence in the consistency of all our foundational modal reasoning, by providing an interpretation on which it comes out true. (This technique resembles the standard method of establishing consistency by building a model. However, in the present case we are not pointing to a collection of objects that form a model but to an intuitively meaningful and non-paradoxical modal notion which ensures that the basic ways reasoning about possibility and necessity used in our justification of the ZFC axioms will be truth preserving.)

However, I should note that Linnebo himself might reasonably not care about leaving the core modal notion used in his potentialist set theory underspecified, (given the different nature of the philosophical projects he's tended to focus on). For Linnebo often seems to be trying to provide a formal system which can be a kind of meeting ground for discussion and comparison between, e.g., classical mathematicians and intuitionists. And if this is what you want to do, then using a notion of $\Box$ and $\Diamond$ in potentialist paraphrases which can be equally naturally given many different philosophical interpretations justifying the acceptance of different inference rules is a benefit rather than a weakness.