Iterative Conceptions of Set

Neil Barton

30. December 2022

Acknowledgements

This short book came to be as a result of a lot of help from a lot of different people. My intellectual debts are great, and any list of thanks is going to be very incomplete. Constraints of space prevent me from going on as long as I would like. But it's still good to try.

I got excellent comments on drafts from Salvatore Florio, Marcus Giaquinto, Sigurd Jorem, Øystein Linnebo, Toby Meadows, Sam Roberts, Chris Scambler, Zeynep Soysal, Sofie Vaas, Mahan Vaz, Pascal Wagner, and Daniel Waxman. I thank all of you heartily for your time, patience, and helpfulness.

In addition, there are many people I have had very useful conversations with regarding ideas behind the text, and whose words helped shape the book in some way. I can remember having especially fruitful interactions with Carolin Antos, Joan Bagaria, Robert Black, Tim Button, Laura Crosilla, Raphaël Carroy, Neil Dewar, Alexander Douglas, Monroe Eskew, Juliet Floyd, Sy-David Friedman, Victoria Gitman, Olav Gjelsvik, Joel-David Hamkins, Luca Incurvati, Akihiro Kanamori, Daniel Kuby, Julie Lauvsland, Maxwell Levine, Penelope Maddy, Beau Madison Mount, Rupert McCallum, Colin McLarty, Michael Potter, Augustín Rayo, Ian Rumfitt, Ethan Russo, Jeffrey Schatz, Stewart Shapiro, Robin Solberg, James Studd, Fenner Tanswell, Philip Welch, and Kameryn J. Williams. My apologies to the large number of people who I have forgotten.

Versions of this work were presented in Oslo and Konstanz, and I thank those audiences for their input.

I would like also to thank Penny Rush and Stewart Shapiro for their editorial work and for bringing together the *Elements in the Philosophy of Mathematics* series. Shalini Bisa and Hilary Gaskin were tremendously patient and helpful with the administrative side of things, and I thank them for their hard work.

I've been very lucky to be part of some great institutions. Birkbeck provided a fantastic place to do graduate work in philosophy. The Kurt Gödel Research Center for Mathematical Logic was a hive of logical activity whilst I was there. Konstanz was both a very supportive and energetic place to do philosophy. And though I haven't been there long, I'm having a great time in Oslo and enjoying myself immensely.

I'm grateful also to the financial support of the UK Arts and Humanities Research Council whilst doing my PhD, the FWF (Austrian Science Fund) during my stay in Vienna, the VolswagenStiftung whilst I was in Konstanz, and Norges Forskningsrådet (Research Council of Norway) who provide my current funding.

I want to also acknowledge the invaluable support of my partner Nikki. She's been a constant source of inspiration, and if the book holds up halfway to the standards she sets for herself, I'll be very happy.

Lastly, I dedicate this book to my parents Jeanne and Paul, who have supported my curiosity for as long as I can remember.

Contents

Chapter 1

Introduction

If you're reading this book then I presume that you're curious about infinity, set theory, and its philosophy. Growing up I'd always been interested in philosophy. Mathematics however, I found to be a necessary but tiresome part of the curriculum, especially through my teenage years. I had great teachers, but the focus on exam preparation that inevitably took up the bulk of our time was just plain boring—solving dreary computational problems using known algorithmic methods (a task that I'm not especially good at to this day). This didn't fit so well with what my mother Jeanne (a mathematics teacher) had always told me—that at a certain point mathematical study can feel like "doors opening left and right". It was at university that I saw Cantor's Theorem and Gödel's Theorems for the first time. Suddenly I understood what my mum had meant—mathematics was an area where new ideas and methods could result in a complete shift in one's perspective on the world, and your ability to solve problems is only bounded by your creativity and the constraints of logical space. The doors were very much open, and I became increasingly interested in notions of infinity in mathematics. To understand infinity, it's very natural to start by considering our best mathematical theories of it. Set theory, as a theory of infinite collections and what we can do with them, was the obvious choice. Understandably, philosophers have showed a lot of interest in set theory since its beginnings in the late 19th and early 20th century. There was already plenty of philosophical material to get my teeth into, and I tucked in with gusto.

What I discovered, however, was that the buffet was far richer than I'd anticipated. In particular, several philosophical and mathematical advances have been made in the philosophy of set theory since the early 2000s. Both mathematicians and philosophers have closely examined ideas concerning whether there is an *all-encompassing domain* for set theory, and how the tools of contemporary set-theoretic practice might bear on philosophy. This has tied the study of the philosophy of set theory very closely to issues in *metaphysics*, including the nature of *possibility* and *absolute generality*. However, I think it's fair to say that these developments (with some notable exceptions) have been passed over for mainstream philosophical consideration. Whilst this is understandable—the mathematical barrier to entry is high and our time is finite—the philosophical issues themselves are (in my opinion) understandable to anyone with some introductory logic courses under their belt.

Many philosophers are aware of the *paradoxes* of set theory (e.g. Russell's paradox). For many people, these were solved by the *iterative conception* of set which holds that sets are formed in stages by collecting sets available at previous stages. This book will examine possibilities for articulating this solution. In particular:

Main Aim. I will argue that there are different kinds of iterative conception, and it's open which of them (if any) is the best.

Along the way, I hope to make some of the underlying mathematical and philosophical ideas behind tricky bits of the philosophy of set theory clear for philosophers more widely, and make their relationship to other questions in philosophy perspicuous.

Here's the plan. Chapter 2 will lay down some reasons as to why we should be interested in set theory as philosophers and mathematicians. This chapter serves as a motivation for the reader less familiar with set theory to get excited, however it also serves a dual purpose—we'll see some desiderata that will be employed later in the book when we come to assess set-theoretic concepts/conceptions.

Chapters 3 and 4 set up a way of thinking of set-theoretic progress as trading off inconsistent principles. Chapter 3 will go over the naive conception of set and the paradoxes that brought it down. We'll also provide a diagnosis of the problem as involving a conflict of two inconsistent principles. This material is well-worn, but I'll explain a twist on the classic paradoxes that has been examined by philosophers recently (namely that we can think of these paradoxes as paradoxes about the existence of *functions*) which will help integrate this material with what comes later. Chapter 4 will present the emergence of the *combinatorial conception* and *logical conception* of set, and give the *iterative conception* as a further sharpening of the combinatorial conception. We'll also explain the standard 'strong' version of the iterative conception, and how it can be given a modal stage theory.

Chapter 5 will then explain some mathematical ideas that have informed the development of contemporary set theory under the iterative conception, namely *forcing* (a way of adding subsets of sets to models). I'll do my best to make these mathematically tricky ideas palatable to philosophers.

Chapter 6 will explain a paradox given forcing for a certain conception of set, and explain how it is linked to the incompatibility of inconsistent principles (much like the naive conception). In particular we'll see how the Powerset Axiom is incompatible with the idea that there should be saturation under forcing in the universe.

Chapters 7 and 8 will identify a split in how we might move forward. Chapter 7 will explain how there is a genuine choice between Powerset and forcing saturation, and will show how forcing saturation can be viewed as arising from kinds of iterative process. Chapter 8 will explain how mathematics is interpreted within each conception, and will contrast each in the light of the theoretical virtues discussed in Chapter 2.

Finally, Chapter 9 will provide a concluding summary and identify some further work that is needed in order to obtain greater clarity on these issues. In particular, I'll explain some salient objections that need addressing in order to move forward. I hope that the reader comes away with a sense of how set theory is philosophically interesting and the vastness of conceptual space.

Before we get going, however, a few remarks are in order. First, whilst I hope that this book is of pedagogical value and can help people new to the philosophy of set theory gain an understanding of some difficult mathematics, *this is not a textbook*. My approach is one of conveying underlying ideas, rather than giving everything in full rigorous detail. Where sensible I've tried to give formal definitions and references for the interested reader in footnotes rather than the body of the text.

Relatedly, the pacing of this book will feel slightly odd. There is a *tension in exposition* in that I both want to get the *newcomer* interested but also accomplish a significant *research-oriented* goal. I therefore run the risk of boring the reader who has been studying these issues for years whilst outstripping what can be expected of an early student (however talented). I've tried to present the known material in such a way that it makes recent novel twists on old material clear, and to keep the harder material as accessible as possible. However this book is *hard* if you aren't familiar with the relevant bits of mathematical logic. My aim is to make things *accessible* and not, per impossibile, *easy*. To combat this problem, the book runs along two tracks. The 'standard' track is intended for those who do not necessarily have years of philosophy of set theory under their belt. The 'expert' track is for those who already know a good bit of philosophy of set theory. I denote sections/paragraphs/footnotes that are on the expert track with a 'blackbelt' emoji 🥋 (and often inside a box). I encourage everyone to read all the book, after all it's helpful to peek behind the curtain and see some of the complicated workings of the machine. But readers should not feel disheartened if 🥋-parts are tricky to follow—those are especially difficult and one shouldn't expect to get everything first try.

I'll use the following conventions. Bits of language (e.g. syntax/utterances) will be enclosed within double quotation marks. So "Toffee is a clever cat" can be a sentence or an utterance, "cat" is a word or term in language, and "Toffee" is a name (in this context), whereas Toffee is a (particular) cat who is also clever. Single quotation marks will be used as 'scare quotes' i.e. cases where the enquoted phrase is not to be taken literally (though it may be illustrative). In cases where such usage occurs in a formal context, single quotes often denote an abbreviation for a formal claim (e.g. $\mathsf{PA} \vdash$ 'There are infinitely many prime numbers', even though "There are infinitely many prime numbers" is a sentence of English, not Peano Arithmetic). Italics are reserved for emphasis, or where they occur in the scope of a definition, the definiendum. I allow definitions to be informal and philosophical as well as formal, but I will clearly separate the informal and formal definitions. With these conventions in hand, let's get ready to set out!

Chapter 2

Why set theory?

Before we start getting into the iterative woods, I want to give some motivation for studying set theory and its philosophy.

Question. Why do this, given that there's so many good introductions into these topics?

Answer. As well as providing a survey of some of the literature, this chapter will lay down some *theoretical virtues* that we might think theories/conceptions of set can have. These virtues will be important later when we come to assessing our options.

What are sets? Here's a rough-and-ready definition:[1]

Definition 1. (Informal) A *set* is a kind of collection that is:

(i) **Extensional:** Sets with different members are non-identical, and sets with the same members are identical.

(ii) **Objectual:** Sets are *objects* over and above their elements.

So, for example, I can consider the set of books currently on my table. This is an object, in addition to the books themselves. If I take a book off my table, the term "the set of books on my table" now denotes a different set, since this new set of books has different members.

Just given this bare bones story, it's natural now to ask: **Why be interested in set theory at all?** It's useful first to consider a *bad* answer (but one that helps us see the *role* of set theory more clearly):

[1]( ) It's plausible that nowadays we think that sets are *combinatorial* too (in the sense of being extensionally equivalent to pluralities of objects, irrespective of whether we can provide a circumscribing definition). Later we'll set up the difference between the logical conception and combinatorial conception of set, and so I don't want to commit to this just yet.

Theory of Collections. Set theory provides our best theory of collections.

This is perhaps encapsulated by George Boolos' claim that:

> I thought that set theory was supposed to be a theory about all, "absolutely" all, the collections that there were and that "set" was synonymous with "collection" [Boolos, 1998, p. 35][2]

The idea that the interest of set theory derives from "set" being synonymous with "collection" or providing our best theory of collections is open to at least two powerful criticisms:

First, there's lots of different ways we talk about collections. To take two simple kinds: (1.) Collection-like talk needn't be **objectual**. As the vast literature on plural logic indicates[3], we can talk about and quantify over objects in the *plural* without thereby committing to a *set* of them. So, instead of talking about the *set* of books on my table, I could just have talked about *the books on my table* in the plural. (2.) Collection-like talk needn't be **extensional**. Instead, it can be taken *intensionally*, where identity is not taken to be governed by an extensionality criterion. Presumably there's a sense in which I don't destroy my *beer coaster collection* just by giving one of the (many) beer coasters to a friend. My collection of beer coasters is just the kind of thing that can survive a loss (or better yet, gain) of some members.

Second, even if set theory did provide our best theory of collections, there's much more to the story. Collections of beer coasters are a perfectly good subject matter for philosophical study, but this observation fails to explain *why* set theory is often regarded as *central* to many areas (and especially mathematics).

Here's what I take to be the core point: *Objectual and extensional collections, when augmented with the 'right' axioms, are powerful devices of representation.* And the ability to *represent* means that all sorts of problems, both philosophical and mathematical, can be encoded within set theory.

Let's look at this idea in a little more detail. This representational power presents two interlinked aspects of set theory:

Foundation for Mathematics. Set theory provides a 'foundation' for mathematics (and hence mathematical tools in philosophy).

Philosophical Repository. Set theory examines many philosophically interesting subjects (e.g. paradoxes, infinity).

[2]() Boolos here is discussing the contrast between sets and proper classes, so perhaps the quotation is intended for a slightly different context. Whatever the weather, the idea that set theory just provides our best theory of collections is enough to get the ball rolling.

[3]See [Florio and Linnebo, 2021] for a book-length treatment.

This division is far from exclusive. Certainly there are cases where we might think that set theory and philosophy are inextricably intertwined.[4] Indeed, this book emphasises the fact that mathematics and philosophy can become fruitfully intermixed, and I do not think it is either necessary or desirable to keep these considerations separate. Nor do I think that *every* bit of set theory will be entangled with philosophy, and there are set theorists who study solely mathematical questions. Still, the distinction serves as a rough categorisation for different facets of set theory.

At this stage, we'll keep things relatively informal, but a little precision will be helpful. One set theory that's proved to be of central interest is *Zermelo-Fraenkel set theory with the Axiom of Choice* (ZFC), which we'll examine more closely later. For now let's just content ourselves with the following rough characterisation: ZFC tells you that there's lots of sets (both finite and infinite) and let's you do many of the usual set-theoretic operations you want on those sets (e.g. take the union of two sets).

Recently, Penelope Maddy has isolated some *mathematical goals* of set-theoretic foundations built on ZFC.[5] I'll provide some examination of Maddy's ideas, and I'll suggest some modifications and additions of my own.[6] These goals serve a dual purpose. On the one hand, they motivate the consideration of set theory for the interested reader. On the other, we will use them later to evaluate particular conceptions of set.

Earlier I mentioned that set theory is a powerful device of *representation*. Many of the desiderata we'll consider are linked to this idea. For instance:

Observation. We can encode/represent all mathematical objects using sets.[7]

What do I mean by 'encode/represent' here? Let's take a simple example from high-school mathematics. We want to consider some geometric object in two-dimensional (Euclidean) space, let's say a straight line. By picking an origin and imposing a coordinate system, we can represent this straight line by some function $f(x) = bx + c$, and think of the straight line as composed of its *graph of ordered pairs* $\langle x, bx + c\rangle$. This can help us when, for example, trying to compute the lengths of line segments (e.g. by using the Pythagorean theorem). But the ordered pairs aren't (intuitively speaking) *the same* as the line, they just *encode* it.

[4]See, for example [Rittberg, 2020] who argues that set-theoretic mathematical practice can be metaphysically laden.

[5]See [Maddy, 2017] and [Maddy, 2019].

[6]For clarity's sake, **Generous Arena**, **Shared Standard**, **Metamathematical Corral**, and **Risk Assessment** are all explicitly identified by Maddy, and **Theory of Collections**, **Foundation for Mathematics**, **Philosophical Repository**, **Theory of Infinity**, **Independence**, **Limits of Thought**, and **Testing Ground for Paradox** are my own additions (though many are implicit in much of the literature and Maddy's work).

[7]See [Posy, 2020], Ch. 2, for a very concise survey of the classical situation (Posy sets up the classical mathematician as a foil for intuitionism), as well as many set theory textbooks (e.g. [Enderton, 1977]).

So with sets, but generalised to any mathematical object you'd care to consider. Zero can be encoded by the empty set, natural numbers by the finite von Neumann ordinals[8], rationals as pairs of natural numbers, reals as Dedekind-cuts of rationals,[9], ordered pairs as Kuratowski-ordered pairs[10], and functions/relations by sets of ordered pairs (i.e. the function f is encoded by $\{\langle x, y\rangle | f(x) = y\}$). Of course there's lots of choices, and this is just an illustration of *one* way you might do things.[11]

Using similar tactics, any mathematical object we have come up with can be encoded by sets (putting aside some controversial cases[12]). This has some important consequences. First, set theory provides a:

Generous Arena. Find *representatives* for our usual mathematical structures (e.g. $\mathbb{N}$, $\mathbb{R}$) using our theory of sets.

I think it is worth pausing for a moment to reflect on just how remarkable **Generous Arena** is. Just using the membership relation and suitable axioms, we can find a representative for almost any object you'd care to discuss—all the vertiginous diversity we see in mathematics can be captured by that one little relation of membership. Because we can encode mathematical objects as sets, we have a way of relating them to each other within a single domain. This Maddy argues, gives us:

Shared Standard. Provide a standard of correctness for proof in mathematics.

The thought here is that because we have **Generous Arena** and can view mathematical objects as encoded/represented by sets, a proof about a mathematical object can be regarded as correct if it could be (in principle) translated into a proof in set theory about properties of the relevant mathematical code(s). Of course, "in principle" is *important* here—outside of set-theoretic mathematics it

[8]These can be defined inductively with $0 =_{df} \emptyset$ and $n + 1 =_{df} n \cup \{n\}$.

[9]A Dedekind cut is a partition of the rational numbers into two non-empty sets A and B, where A is closed downwards and does not contain a greatest element.

[10]The *Kuratowski ordered pair* is given by $\langle a, b\rangle =_{df} \{\{a\}, \{a, b\}\}$.

[11]See [Barton et al., 2022] for some of the formal details and further citations.

[12]() For example, one controversial objection (e.g. [Mac Lane, 1986], [Muller, 2001]) to set theory goes something like this: "Everything in set theory has to be encoded by a set, and we know that some categories like the category of all sets are too big to be encoded by sets. So set theory cannot provide a foundation for category theory." I do not find this objection convincing for the following two reasons. (1.) set theorists certainly *seem* to talk about proper-class-sized objects—the study of proper classes is in my (controversial) opinion a perfectly legitimate part of set theory, and (2.) I don't think that category-theoretic study of the sets is really directed at the study of *all the sets*, but rather the study of the schematic first-order properties that all the sets happen to satisfy. A full defence of this idea will have to be left for a different day, but a more detailed explanation of this point can be found in [Barton and Friedman, 2019] (esp. §10.3).

is very clunky to work with these codes, and we shouldn't expect mathematicians to actually go about their daily lives solely using the language of set theory. The relevant language of the discipline in question is probably more flexible than working with just membership. (A desire for a foundation that "will capture the fundamental character of mathematics as it's actually done, that will guide mathematicians toward the truly important concepts and structures, without getting bogged down in irrelevant details" Maddy terms **Essential Guidance**, and since all set theories we'll consider here perform pretty badly in this respect, we'll set it to one side.)[13]

The ability to manipulate large infinite collections in ZFC-based set theory yields the following:

Theory of Infinity. Set theory provides our best theory of infinite numbers.

Theory of Infinity will be important later and so I've explicitly identified it as a theoretical virtue in contrast to some of the literature that leaves it implicit (it does not occur, for example, amongst the virtues identified by [Maddy, 2017] and [Maddy, 2019]). To see its significance, we start by examining the two main kinds of infinite number in set theory, namely *ordinal* and *cardinal* numbers. An *ordinal* number can be thought of as an answer to the question of how *long* an infinite ordering is. Call a set x (under a linear relation R) *well-ordered by* R iff every subset of x has an R-least element. If x is well-ordered by R, then there's no way of descending infinitely in x along R. This helps us think of performing *actions* or *operations* into the infinite along a suitable infinite relations. Within ZFC one can represent and develop an arithmetic for these orders, defining notions of *ordinal addition*, *multiplication*, and *exponentiation*.[14] This provides us with ways of generalising normally finite operations (e.g. computation) into the infinite. [15]

Cardinal numbers, by contrast, can be thought of as answers to the question of how *many* objects there are in a set. In particular, we say that two sets X and Y *have the same cardinality* iff there is a bijection between them, where a *bijection* $f : X \rightarrowtail\!\!\!\!\rightarrow Y$ is a function that 'pairs off' the members of X and Y, i.e. f takes no two elements of X to the same element of Y (f is *injective*) and every element of Y is hit by f applied to some element of X (f is *surjective*). By representing cardinals using particular kinds of sets, ZFC provides a theory in which the cardinal sizes of any sets can be compared and natural operations like

[13] See [Maddy, 2017, p. 305].

[14] There's lots of ways to do this, but one popular way is to use von Neumann ordinals, where we let $0 = \emptyset$, $\alpha + 1 = \alpha \cup \{\alpha\}$, and limit $\lambda = \bigcup_{\beta<\lambda} \beta$. Addition is represented by the *ordered disjoint union*, multiplication by the *lexicographical ordering* on the *product*, and exponentiation by *iterated multiplication*.

[15] ( ) For example, we can consider *infinite time* Turing machines. See [Hamkins and Lewis, 2000].

multiplication, addition, and exponentiation generalised and computed.[16] The success of ZFC is striking, it seemingly gives finite beings (e.g. us) the ability to reason about large infinite objects. Many surprising facts can be thereby shown. For example, we can prove that:

Theorem 2. There are as many natural numbers as there are squares of natural numbers (in particular $f(x) = x^2$ is just such a bijection from the natural numbers to the squares of naturals).

This is somewhat surprising since the squares of n and $n + 1$ get more and more spread out as n gets larger. Indeed, similar results were even regarded as kinds of 'paradox' by Thābit ibn Qurra and Galileo. We can even show:

Theorem 3. The set of all *rational numbers*—the numbers expressible by fractions—is the same size as the set of all natural numbers.[17]

This is so even though there are infinitely many rational numbers between any two natural numbers. We can also show:

Theorem 4. There are as many real numbers between 0 and 1 (or any two real numbers for that matter) as there are in the real line, or in any n-dimensional plane based on the real line (i.e. $\mathbb{R}^n$).[18]

Despite these surprising results on *sameness of size*, we also discovered that infinity comes in *different* cardinal sizes:

Theorem 5. (*Cantor's Theorem for the reals*) The cardinality of real numbers is greater in size than the cardinality of the natural numbers, in the sense that (i) there is no bijection between the natural numbers and the real numbers, and (ii) there is an injection from the natural numbers to the real numbers.

This phenomenon appears to be more general than merely comparing the natural numbers and real numbers. We in fact discovered that:

Theorem 6. (*Cantor's Theorem*) Let $\mathcal{P}(x)$ denote the *power set* of x, the set of all subsets of x (that such a set always exists is one of the central axioms of ZFC). Then the cardinality of $\mathcal{P}(x)$ is *greater than* that of x.[19]

---

[16] Again, there's a variety of ways one might proceed, but here's a typical one. The *cardinality of X* can be represented as the least von Neumann ordinal bijective with X. Cardinal addition can be computed as the cardinality of the disjoint union, multiplication as the cardinality of the product, and exponentiation X^Y as the cardinality of the set of all functions from Y to X.

[17] See, for example, Chapter 2 of [Giaquinto, 2002] for an explanation of this result.

[18] Again, see Chapter 2 of [Giaquinto, 2002].

[19] We'll discuss a proof of Cantor's Theorem later, in particular as it relates to the paradoxes in Chapter 3.

Again, Cantor's Theorem is striking. It seems to imply, on the basis of natural principles about sets, that if there's one infinite set then there's a *never ending hierarchy* of infinite sets, since the power set of any set x is always bigger than x. Moreover, it produces much of the interest of cardinal arithmetic—whilst addition and multiplication are trivial for infinite cardinal numbers (one can show that both addition and multiplication just result in getting the larger of the two back) cardinal exponentiation is *not*—one can show that $2^\kappa > \kappa$ for any cardinal κ.[20]

The ability to work with infinity plays out in various areas of philosophy, including areas outside the philosophy of mathematics.[21] Indeed, these arguments are often regarded as a refutation of the time-honoured position in philosophy and mathematics that infinity is completely beyond understanding and intractable within mathematics.[22]

However, this success must be tempered by the following phenomenon that emerged in the 20th century:

Independence. There are sentences of set theory that can neither be proved nor refuted using our 'canonical' theory of sets ZFC, assuming that ZFC is consistent. Nor can any 'reasonable' expansion of ZFC settle all questions formalisable in the language of set theory.[23]

Before we discuss this further let's remark that the *mere fact* of independence

---

[20]In particular, you can think of 2^κ as the size of $\mathcal{P}(\kappa)$, since any member of $\mathcal{P}(\kappa)$ can be correlated with a unique function from κ to $2 = \{0, 1\}$ via *characteristic functions* (where for $X \subseteq \kappa$, $f(\alpha) = 1$ iff $\alpha \in X$).

[21]() Here's an example from infinite ethics showing how infinite assumptions can play out with utility calculations (the example is due to [Cain, 1995]). Suppose we have people arranged at all coordinates of the real plane indexed by integers (so there's a single person at every (m, n) for integers m and n). A circle slowly grows from the origin. In one scenario (the *circle of happiness*), everyone starts at utility -1 and moves to utility $+1000$ (or any large finite amount) when they fall inside the perimeter of the circle (and remains at this value forevermore). For the *circle of negativity*, each agent starts at $+1$ and goes to -1000 when they get caught by the circle. With simple cardinality arguments one can argue that the sum of the utility for the expanding sphere of negativity is positively infinite, whereas the expanding sphere of happiness is negatively infinite (one needs to define these terms, but the rough idea is that there's always boundedly many happy/sad people in the circle of happiness/negativity, whereas infinitely many people of the opposite disposition). Cain argues that we should nonetheless *prefer* to be in the expanding happiness world (since then we just have to wait long enough to be blissfully happy forevermore). Thanks to Joel David Hamkins for communicating this example to me, see [Hamkins and Montero, 2000] for some further discussion.

[22]See, for example, the paradoxes of the infinite given in the Introduction to [Moore, 1990]. The place of Cantor, his results, and other scholars in arriving at a final acceptance of infinity is actually somewhat more subtle than is often acknowledged (see [Ferreirós, 2007], especially the Introduction). () In particular, it is somewhat unclear whether our notion of cardinality *had* to be the Cantorian one, or if in different circumstances we might have ended up with a version of cardinality that respects the idea that a proper part should always be smaller than the whole.

[23]() Here 'reasonable' means recursively enumerable and consistent.

is philosophically important. It shows that there will be limits to what our formal theories capture. There are at least two kinds of independence that will be relevant for us. To set things up, let's start with the following:

Definition 7. We let the cardinal numbers be indexed by ordinals using a function we'll call the 'aleph' function (or $\aleph$). $\aleph_0$ is the smallest cardinal number (which happens to be the cardinality of the natural numbers). $\aleph_1$ is the next smallest, and more generally $\aleph_\alpha$ is the α^{th} cardinal number. We'll denote the ordinal corresponding to $\aleph_\alpha$ by ω_α (we'll often also let ω_0 denoted by "ω").

A routine argument shows that $2^{\aleph_0} > \aleph_0$ (by Cantor's Theorem). But is there anything in between? That is, does $2^{\aleph_0} = \aleph_1$? Or are there cardinalities in between, and in fact $2^{\aleph_0} > \aleph_1$?

Definition 8. We will use the following for discussing the spread of cardinalities:

- The *Continuum Hypothesis* (or CH) is the statement that $2^{\aleph_0} = \aleph_1$.
- The *Generalised Continuum Hypothesis* (or GCH) is the statement that 'For every ordinal α, $2^{\aleph_\alpha} = \aleph_{\alpha+1}$' (i.e. every jump in cardinality obtained by applying the powerset operation to an infinite set just pushes you up *one* cardinal number).
- The *continuum function* is defined by $f(\aleph_\alpha) = 2^{\aleph_\alpha}$ (i.e. the function that takes an infinite cardinal to the cardinality of its powerset).

As it turns out, CH, $\neg\mathsf{CH}$, GCH, and $\neg\mathsf{GCH}$ are all consistent with ZFC (assuming ZFC itself is consistent). We'll explain how this works later (Chapter 5).

To discuss the other kind of independence, we first need a brief foray into *consistency strengths*. Within arithmetic, and hence within ZFC, one can (computably) encode syntactic notions like *sentence*, *formula*, *proof*, and *consistency*. This allows you to formulate a sentence within ZFC expressing the idea that ZFC is itself consistent (more precisely, you can formalise within ZFC the sentence that there's no proof of a contradiction derivable from the axioms of ZFC). Call this sentence $Con(\mathsf{ZFC})$. But now we can point to:

Theorem 9. (*Gödel's Second Incompleteness Theorem*) Assuming that ZFC is consistent[24], then $Con(\mathsf{ZFC})$ is not provable within ZFC, and nor is $\neg Con(\mathsf{ZFC})$. Moreover, this theorem holds for any (suitably nice[25]) theory that can represent arithmetic.

[24] Strictly speaking we need ω-consistency, and this is Rosser's strengthening, but we'll put this to one side for the sake of clarity.

[25] i.e. recursively enumerable and consistent.

Within set theory we can study a wide variety of sentences that have different consistency strengths—one can often prove one extension of ZFC consistent from another. As it turns out, CH and ¬CH are *not* like this (ZFC, ZFC + CH, and ZFC + ¬CH are all *equiconsistent* in that one can prove each consistent from the other). Obviously adding $Con(\mathsf{ZFC})$ results in a consistency strength increase. There are other principles—so called *large cardinal axioms*—that are important here. These serve as the natural indices for consistency strength. They postulate the existence of sets with a lot of *closure* properties and if they exist (or are consistent) we can prove that many theories are consistent by finding *models* of the relevant kind. Set theory has in fact discovered a whole hierarchy of these cardinals with stronger and stronger closure properties.

() Here's an example:

Definition 10. A cardinal κ is *strongly inaccessible* (or just *inaccessible*) iff:

(i) κ is uncountable (i.e. it's bigger than the cardinality of natural numbers).

(ii) Given any set x smaller than κ, the cardinality of $\mathcal{P}(x)$ is also smaller than κ (in this case we call κ a *strong limit cardinal*).

(iii) Given any set x smaller than κ, and any function $f : x \rightarrow \kappa$, the range of f is bounded by some $\gamma < \kappa$ (here, we say that κ is *regular*).

It's instructive to think about what such an axiom says. Such a κ seems very big—clause (i) ensures it's bigger than $\mathbb{N}$, (ii) says that you can't catch it with something smaller by taking our favourite size-increasing operation (powerset), and clause (iii) says that you can't catch it by mapping a smaller object into it using a function. One can show that an inaccessible cardinal κ suffices to produce a model for ZFC (and much more), and so by Gödel's Second Incompleteness Theorem you can't produce an inaccessible cardinal from ZFC alone. We can strengthen this axiom by postulating that there is a cardinal κ that is (i) strongly inaccessible, and (ii) has κ-many strongly inaccessibles beneath it. And these cardinals lie *right at the bottom* of the large cardinal hierarchy.[a]

[a] See, for example, the diagram on p. 472 of [Kanamori, 2009] for an idea of the extent of the space.

Those are the two kinds of independence we'll consider. One (the CH kind) involves the exact value of cardinal sizes. The other (the large cardinal kind) involves considering sets with ever greater and greater closure properties and consistency strength. These aren't the only kinds of independence (there are also strong axioms that don't directly postulate the existence of large cardinals[26]) but these are the ones we'll focus on.

We should pause for a moment to reflect on what this independence tells us about our ability to provide formalisations of theories of sets. Whilst ZFC does give us the resources to prove a great many things about the infinite, it does not yield information about the values of many cardinal computations nor what kinds of set exist with certain closure properties. How we might respond to this situation will be a central theme of this book, but it should be noted that **Independence** is a reason for philosophers—i.e. not just mathematicians—to be interested in set theory. Assessing the impact of independence is central for understanding our place in the world and what we can (and maybe can't) do. I think it's important therefore to isolate the following philosophical aspect of set theory.

Limits of Formalisation. Set theory provides a natural place to examine the limits of our formalisation, pushing the boundaries of what might be realistically expected to be captured, and exploring where formalisations may finally give out.

It's a beguiling question to think what the implications of **Limits of Formalisation** might be. Does it imply that there are limits on what can be known? Or that there is some kind of metaphysical indeterminacy in the world? These are important questions for philosophers, and show that **Independence** is more than a mere mathematical curio.

From the mathematical perspective, set theory is one of the main theories in which we study **Independence**. It provides us with flexible tools with which we can study models of different theories, how they can be built from one another, and hence how relative provability works (given the Completeness Theorem). We can thus (with Maddy) identify:

Metamathematical Corral. Provide a theory in which metamathematical investigations of relative provability and consistency strengths can be easily conducted.[27]

As philosophers, we should be keen to assess whether the theories we work in are consistent. **Metamathematical Corral** combined with the fact (as we'll

[26] See, for example, so called 'Axioms of Definable Determinacy' [Koellner, 2014].

[27] As experts will know there are other theories we might pick. One only really *needs* a theory of syntax to study consistency (and weak theories of arithmetic suffice for such a theory). Another salient field here is proof theory and the study of proof-theoretic ordinals. In a way, set theory provides *more* than what is required for examining **Metamathematical Corral**. However, it is in the variety of *models*, and what one can build from them, where set theory really shines. So it is perhaps better to say that set theory provides a piece of the puzzle for **Metamathematical Corral**, rather than the whole picture. Thanks to Marcus Giaquinto and Daniel Waxman for some further discussion here.

see later) that set theory often comes with an attendant conception of what the sets are like gives us:

Risk Assessment. Provide a degree of confidence in theories commensurate with their consistency strength.

In particular, suppose that you come up with a wild new theory T (either philosophical or mathematical). If I can use some set theory S to produce a model of T, then I know that I can be at least as confident in the consistency of T as I am in S.

Risk Assessment is especially important as many theories here are *in*consistent. As many philosophers know, early set theory was subject to paradoxes (e.g. Russell's Paradox). However set theory can also yield inconsistency and paradox when combined with other philosophical principles, such as when we layer mereology on top of the sets (e.g. [Uzquiano, 2006]). I also want to point out (in line with **Philosophical Repository**) that an *enormous* variety of set-theoretic ideas can be extended to inconsistency. In particular when we push ideas to their natural limit, they nearly always explode, Perhaps this constitutes a kind of 'paradox' (maybe in a weak sense of the term). Some of these we'll see later, and some others I mention in a footnote for the reader who wants to look further.[28] One might think that this is a negative of the discipline—after all isn't inconsistency a (if not *the*) unforgivable sin? I disagree. Inconsistency can be informative. Set theory gives us the tools to locate and diagnose these inconsistencies, helping us to elucidate our **Limits of Formalisation** and further giving us a:

Testing Ground for Paradox. Set theory is very *paradox* prone, both in terms of the principles that can be formulated within set theory and when combined with certain philosophical ideas (e.g. absolute generality and mereology). In this way, set theory provides a *testing ground* for seeing when and how ideas are inconsistent.

So, there's some interesting and nice features of set theory—not just a **Theory of Collections**, but a field that provides a **Foundation for Mathematics** and **Philosophical Repository**, in particular by providing a **Generous Arena**, **Shared Standard**, **Theory of Infinity**, the example of **Independence** and its use as a **Testing Ground for Paradox**, that help articulate the **Limits**

[28] For example, the embedding template $j : V \to M$ for large cardinals explodes when $M = V$. Forcing axioms can pop in various ways, either by admitting too many parameters, allowing too many kinds of forcing, or not keeping a tight enough control on the sentences allowed (see [Bagaria, 2005]). Standard reflection principles blow up at the level of third-order reflection (see [Reinhardt, 1974] and [Koellner, 2009]) and modal reflection principles are pretty flammable too (see [Roberts, 2019]).

of Thought, give us a **Metamathematical Corral**, and **Risk Assessment** for our theories. Before we move on, I want to identify one last important aspect of set theory. Although many of these above constraints are simply reasons to be interested in set theory, or were things that set theory happened to be useful for, there is a sense in which set theory was *designed* to fit these purposes. **Risk Assessment**, for example, can't go ahead without set theorists *deliberately* studying **Independence** and **Metamathematical Corral**. In this way, many of the above—notably **Generous Arena**, **Shared Standard**, **Theory of Infinity**, **Metamathematical Corral**, and **Risk Assessment**—are not just pleasant features of set theory, but constraints/desiderata on its development too. Indeed this is one of the central points of [Maddy, 2017] and [Maddy, 2019] (though she leaves **Theory of Infinity** implicit). Thinking about these virtues in this dual light will help to illuminate some of the issues later, and in particular whether different conceptions/theories of sets are *virtuous*.

Chapter 3

The naive conception of set and the classic paradoxes

We've now got some desiderata for set theory on the table (Chapter 2). In this chapter I want to explain one role for conceptions of set (namely to motivate theories) and revisit some well-known material on the naive conception of set and the 'classic' set-theoretic paradoxes. In doing so, I'll present a way of looking at the paradoxes in terms of *functions*.

3.1 Conceptions of set and motivating theories

One way into the problem of the paradoxes is by considering the following:

Question. What do we *want* out of a conception of set?

At least in a mathematical context, what we want out of a conception is a motivation for a *theory*, in particular an *axiomatic theory*.[1] I'll assume that the reader has some understanding of formal axiomatic theories (later we'll use a little bit of first-order predicate logic, plural logic, modal logic, and set theory). But where possible, I'll provide informal paraphrases and reference away the formal details. Further (as outlined in Chapter 2) we want a theory that can do various foundational jobs for us. As good philosophers, it's natural to want conceptual underpinnings, and in particular a conception of set that delivers a theory with the requisite features. This, at least in this book, is what I'll take the primary role of a conception to be—to provide a story of what the sets are like, in order to motivate a particular axiomatic theory of sets. This motivation might take the form of a formalisation (as we'll see later with various *stage theories*), but equally it could be something more informal.

[1] Here I am following some of the remarks in Ch. 1 of [Incurvati, 2020].

3.2 The naive conception of set

Our first such conception will be the naive conception of set:

Definition 11. (Informal) The «*naive conception*» of set holds that sets are extensions of predicates, where the extension of a predicate is the collection of all the things to which the predicate applies.[2]

We now want to consider what axioms the naive conception motivates. For this, it will be helpful to set up an important language for us:

Definition 12. The *language of set theory* or $\mathscr{L}_{\in}$ is the first-order language with one non-logical binary predicate "$\in$" and well-formed formulas formed in the obvious way.

The naive conception clearly motivates adoption of the *extensionality axiom* (which says that any two sets with the same members are equal) as it a conception of set. Unfortunately, it also motivates:

Definition 13. The *Naive Comprehension Schema* asserts that for every one place formula $\phi(x)$ in the language of set theory $\mathscr{L}_{\in}$, there is a set of all and only the sets satisfying $\phi(x)$. Formally:

$$(\exists y)(\forall z)\big(z \in y \leftrightarrow \phi(z)\big)$$

Sadly, as we know, the Naive Comprehension Schema is inconsistent. Let's see how.

3.3 The paradoxes

Why go over the paradoxes, when excellent introductions are available in a wide variety of texts?[3] Aren't I just rehashing old material? Here's why we'll look at them:

(1.) Part of what we will see later is a 'new' kind of paradox (the Cohen-Scott Paradox) and we'll discuss how it's similar to the classic paradoxes. So getting them on the table early is a good idea.

(2.) There has been a shift of focus in the philosophical literature that will help us to see the force of some problems later. Importantly, each of the paradoxes can be linked to the (non-)existence of particular *functions*.

[2] This formulation is taken directly from [Incurvati, 2020, p. 24].

[3] See, for example, [Giaquinto, 2002], [Potter, 2004], and [Incurvati, 2020], for philosophical introductions to the paradoxes, but almost any introductory text on set theory will cover them.

In this book, I'll only really consider Russell's Paradox and Cantor's Paradox. The Burali-Forti Paradox is also interesting, however it is complicated by the fact that one has to use set-theoretic codes for the ordinals (which otherwise could be thought of as sui generis mathematical objects).[4] Here they are:

Russell's Paradox. Consider the condition $x \notin x$. By Naive Comprehension, this determines a set r. We ask: "Is $r \in r$?" If yes, then $r \notin r$ (since r is in the set of all $x \notin x$), contradiction. So instead assume $r \notin r$. Then r satisfies the condition $x \notin x$, and so $r \in r$, contradiction. But then $r \in r \leftrightarrow r \notin r$, a contradiction!

Cantor's Paradox. Consider the condition $x = x$. Let $\{x|x = x\}$ be denoted by u (for "universal set"). Now consider $\mathcal{P}(u)$, i.e. the *power set* of u. By Naive Comprehension, this is also a set. Now we show $x = \mathcal{P}(u)$ by noting: (i) every element of $\mathcal{P}(u)$ is an element of u (trivially), and (ii) if $x \in u$ then $x \in \mathcal{P}(u)$ (since if $x \in u$, then $\forall y \in x, y \in u$ (i.e. $x \subseteq u$) and so $x \in \mathcal{P}(u)$). So $u = \mathcal{P}(u)$.

Clearly then, there is a surjection $f : u \twoheadrightarrow \mathcal{P}(u)$. Now consider the set $c = \{x|x \notin f(x)\}$. Since f is surjective, there is a $y \in u$ such that $f(y) = c$. We now ask "Is $y \in c$?" If yes (i.e. $y \in c$), then $y \in f(y)$, but then y violates c's defining condition, and so $y \notin c$, contradiction. So then we assume $y \notin c$. But then $y \notin f(y)$, and so y meets c's defining condition, and $y \in c$, contradiction. So $y \in c \leftrightarrow y \notin c$, a contradiction!

In fact, this proof can be transformed into a proof of Cantor's *Theorem*, just by replacing u by any old set x and performing a reductio on the claim that there is a surjection $f : x \twoheadrightarrow \mathcal{P}(x)$.

So far, so well-known. Many introductory textbooks contain a presentation of the paradoxes. However, something philosophers have paid more attention to recently (though has been known for a long time) is that these paradoxes are *closely related*:[5]

The Cantor-Russell Paradox. Define u and $\mathcal{P}(u)$ as in Cantor's Paradox. Now consider the case where our surjection $f : u \twoheadrightarrow \mathcal{P}(u)$ is the *identity map* $f(x) = x$. Now the problematic set $c = \{y|y \notin f(y)\} = \{y|y \notin y\} = r$. We'll also refer to this a the Cantor-Russell reasoning.

[4]For some discussion of these issues, see [Menzel, 1986], [Shapiro and Wright, 2006], [Menzel, 2014], [Barton, 2021], and [Antos et al., 2021].

[5]See in particular, [Bell, 2014], [Whittle, 2015], [Meadows, 2015], [Whittle, 2018], [Incurvati, 2020], [Scambler, 2021], and [Builes and Wilson, 2022].

The important thing to note is that in this context (where f is the identity map) the contradictory set r we get out is the problematic set for *both* the Cantor and Russell reasoning (since f is the identity map here, the set $\{y|y \notin f(y)\}$ *just is* $\{y|y \notin y\}$). So the two are not just *superficially* similar, but in many contexts come down to definition of *exactly the same set*, and the core issue is whether there's a surjection $f : u \twoheadrightarrow \mathcal{P}(u)$.

This observation works in the other direction too, where we assume that we have an *injection* (i.e. one-to-one function) $f : \mathcal{P}(u) \rightarrowtail u$. Without loss of generality, again this can be the identity map (since $\mathcal{P}(u) = u$). Now we can just consider the set $\{y|y \notin f^{-1}(y)\}$ (this is well defined since f is an injection).

Cantor's Paradox and Russell's Paradox might still not be *exactly* the same (Cantor's Paradox uses a bit more machinery than Russell's, e.g. injections), but there are clearly strong similarities between the two. I'll remain neutral on whether they are really 'the same' in any deep sense. Important for later will just be:

(1.) We can view each paradox as starting by postulating the existence of a particular kind of function (either a surjection or an injection).

(2.) We can then identify sets x and y such that $x \in y \leftrightarrow x \notin y$ (in the case of Russell-Cantor, x and y are both r).[6]

3.4 Diagnosis

So, Naive Comprehension leads to contradiction. But *why*, and what *options* are we left with? Many have been considered throughout the literature, surveys are available in [Giaquinto, 2002], [Priest, 2002], and [Incurvati, 2020]. We'll follow Incurvati's presentation here, since it will be instructive for making comparisons.

Let's start by noting that the Naive Comprehension Schema encodes the following principle about the concept of set:

Universality. A concept/conception C is universal iff there exists a set of all the things falling under C.[7]

Universality clearly follows from the naive conception, since the condition $x = x$ is a perfectly legitimate predicate of set theory and the naive conception immediately licences the Naive Comprehension Schema. However, the following is also a consequence:

[6] Of course, strictly speaking, anything follows from the contradiction in classical logic. The point is just that a natural way of reasoning to the contradiction is to note the contradictory membership conditions.

[7] This is adapted from [Incurvati, 2020, p. 27].

Indefinite extensibility. A concept/conception C is indefinitely extensible iff whenever we succeed in defining a set u of objects falling under C, there is an operation which, given u, produces an object falling under C but not belonging to u.[8]

Indefinite extensibility also follows from the Naive Comprehension Schema/ This is because any time we have a set x, the Naive Comprehension Schema gives us the juice required for the Cantor-Russell reasoning, and we can then diagonalise to find a set not in x (e.g. one of the members of $\mathcal{P}(x)$).[9]

Clearly, any conception that validates both **Universality** and **Indefinite Extensibility** will be inconsistent, since there both must and can't be a set of all objects falling under the conception. So in order to proceed, a natural way to go is to examine conceptions of set that drop one of these fundamental principles. And this is just what iterative set theories do.

[8]Again, adapted from [Incurvati, 2020, p. 27].

[9]This way of looking at things has clear affinities with [Priest, 2002]'s characterisation of the Inclosure Schema and Domain Principle. Since we're focussed on set theory here, and Priest's framework is more general, I've chosen to go the Incurvati-route.

Chapter 4

The weak and strong iterative conceptions

We found ourselves in a tricky situation at the turn of the 19th century. The burgeoning field of set theory was clearly *useful*, but the naive conception of set was *deeply* flawed. In this chapter, I want to present the emergence of several conceptions of set and the eventual rise of the strong iterative conception. We'll see that this idea can be formalised *modally* and there's a *close affinity* with ZFC. I also want to indicate that there were other conceptions on the market. Analysis of some of the history here in terms of conceptions helps to make sense of the intellectual landscape various agents have inhabited. In particular we can better see that there were possible places where our conception may have diverged.

4.1 The logical and combinatorial conceptions of set

Earlier, we remarked that we want a conception of set to motivate a useful theory of sets. Given the background of classical logic, inconsistent theories of sets are *trivial* (everything follows by the principle of explosion). In motivating a good theory, it's thus preferable if a conception can explain *why* the paradoxical collections don't exist. So a conception should:

(1.) Provide a reasonably 'natural' picture of what the sets are like.

(2.) Motivate a nice theory that is (hopefully) consistent.

In this latter regard, we also need to provide a:

(3.) **Paradox Diagnosis.** Explain why the paradoxical collections aren't sets and which conditions determine sets (and which don't).[1]

[1] Thanks to Sam Roberts for some discussion of this issue.

Let's start to consider some candidates.

Definition 14. (Informal) The *logical conception* of set holds that sets are the extensions of *well-defined predicates*.

The logical conception is often taken to contrast sharply with:

Definition 15. (Informal) The *combinatorial conception* of set holds that sets are extensions of *available pluralities*.

We should note that whilst these conceptions are pretty rough and ready, each admits of multiple different sharpenings (we'll say a little more about this in a second). In this sense, the conceptions are still defective in that they are not sufficiently informative to motivate a good formal theory. However, each indicates the shape of a response to the paradoxes. Many ways of making the logical conception precise will hold that the predicate $x = x$ is 'well-defined'.[2] Thus, under the logical conception, **Universality** is likely to be validated and **Indefinite Extensibility** violated. Conversely, versions of the combinatorial conception will make the notion of 'availability' precise in a variety of ways, and doing so can make it the case that not all sets are 'available'. So it is **Universality** and not **Indefinite Extensibility**, that is likely to be the culprit.

This book does not aim to articulate fully all the different ways we might sharpen the logical conception and combinatorial conception. The reader is directed to [Incurvati, 2020] and the literature contained therein for a much more thorough discussion. However, some things can be said about how one might make each precise. First, we should note that this distinction need not be sharp (a fact that we will return to in §9.5). Still, in order to improve the logical conception, we need to say what "well-defined" means. There are a number of ways of doing this. One (the stratified conception) holds that there are certain formulas that are appropriately stratified, and that comprehension should be restricted to these formulas.[3] Another (the *iterative property conception*) holds that

[2] [Incurvati, 2020] disagrees with this, holding that the limitation of size conception is logical but does not validate **Universality**. It's not clear to me that the limitation of size conception is in fact logical, but in any case we can view those (more precise) conceptions that validate **Universality** as the relevant contrast cases for what we're doing here. Later (§9.5) we'll argue that the classifications amongst the conceptions might not be sharp.

[3] ( ) The stratified conception is proposed by [Quine, 1937] and its history is nicely outlined in [Incurvati, 2020]. One starts with the following definition:

Definition 16. A formula ϕ in the language of set theory is *stratified* iff there is an assignment of natural numbers to variables such that:

(i) For any subformula of ϕ the form $x = y$, the natural number assigned to x is the same as the number assigned to y.

(ii) For any subformula of ϕ the form $x \in y$, the natural number assigned to y is one greater than the number assigned to x.

there is a way of iteratively individuating those formulas that can be used in comprehension (this is the approach of the property theories of [Fine, 2005], [Linnebo, 2006], and [Roberts, MSa]). On each, the predicate $x = x$ is well-defined and individuates an extension.

The combinatorial conception, by contrast, needs to make precise what "available" means. One way is to say that some sets are available iff they can be depicted as part of a particular kind of graph (the *graph conception*). This conception, as it happens, also validates **Indefinite Extensibility** but refutes **Universality**.[4] The *iterative conception* (the focus of this book) also refutes **Universality** whilst accepting **Indefinite Extensibility** (as we'll see in more detail shortly). So each conception is going to either deny **Universality** and accept **Indefinite Extensibility** (as many combinatorial conceptions seek to do) or accept **Universality** and deny **Indefinite Extensibility** (in the manner of many logical conceptions). This division is perhaps not quite perfect (it is unclear if *all* versions of the logical conception and combinatorial conception conform to this template) but it is an instructive way of thinking about set-theoretic progress for what we'll do here.

With this in mind, let's start to see how this plays out with respect to the focus of this book, namely 'the' iterative conception of set.

4.2 Some iterative conceptions of set

We'll consider a kind of combinatorial conception known as the iterative conception. We'll keep things rough and imprecise to begin with (this imprecision will be helpful later when we separate out different versions of it):

Definition 17. (Informal) The *iterative conception* of set holds that sets are formed in stages, and new sets are formed from old by collecting together sets formed at previous stages. There are no other sets than those found at the stages.

The rough idea can be filled out as follows. We (or better—a suitably idealised being) start at an initial stage with some initially given collection of objects. These could be a bunch of non-sets (often called *Urelemente*), or some antecedently given sets that we take to be acceptable (e.g. the empty set).[5] We then

By restricting comprehension to stratified formulas we obtain a system known as NF.

[4] Since we won't discuss this much, we'll set it aside, but see [Incurvati, 2020], Ch. 7 for details. () The relevant notion is an *accessible pointed graph*, a kind of directed graph where there's a distinguished top node (this is the 'pointed' part of the definition, and you can think of this 'point' as the set we want to code), the edges code the membership relations, with accessibility meaning that it's possible to reach each node of the graph by some finite chain of edges starting from the point.

[5] () Depending on what set-forming operations we allow, we have to be careful that we don't start with a proper class. If we do, some modification is needed, see for example [Menzel, 1986] and [Menzel, 2014].

begin forming new sets out of what we have using some given operations, and in this way obtain the sets. So long as our operations always guarantee that new sets can always be formed, we have an explanation of why **Indefinite Extensibility** holds and **Universality** fails—there will never be a stage at which we can use an operation to collect all the sets into a set.

The iterative conception of set as I've given it can in fact be split into two conceptions, a strong one and a weak one:[6]

Definition 18. (Informal) The *strong iterative conception* of set holds that sets are obtained in a sequence of stages. At each additional stage we form *all possible subsets* of sets available at previous stages. There are no other sets beyond those obtained this way.

Definition 19. (Informal) The *weak iterative conception* of set holds that sets are formed in stages. Sets are formed by collecting together sets at previous stages using some set-forming operations. We leave it open whether or not we get every possible subset of what we have at a stage immediately after the current one. There are no other sets beyond those obtained this way.

I want to suggest that the weak iterative conception is really *prior* to the strong iterative conception (conceptually, if not chronologically). Key to the weak iterative conception are:

(i) A description of what counts as a starting domain.

(ii) A description of some operation(s) for forming new sets from old.

The strong iterative conception says (i) can be any set of objects, but the empty set will do, and (ii) that the operations consist solely of powerset (i.e. taking all possible subsets). It thus *sharpens* the weak iterative conception; since there are other set-forming operations that we might have chosen. Let's see an example of the difference by going into more detail on each.

The strong iterative conception is perhaps the simplest version of the weak iterative conception, so we'll explore it first. It is also perhaps the 'default' version—as of writing, if you put the terms "iterative conception of set" into a search engine, you'll get back results about the strong iterative conception. Assuming that we think that any subset of a set could exist, the strong iterative conception amounts to asserting that the sets are obtained by iterating the powerset operation (and if we have an infinite chain of stages we can throw them together into a new stage). This idea can be formalised with the following definition using ordinal numbers:

[6] This distinction emerged in discussion with Chris Scambler, and I'm grateful to him for the suggestion of separating out the two.

Definition 20. *The Cumulative Hierarchy of Sets* or V is defined as follows:[7]

(i) $V_0 = \emptyset$

(ii) $V_{\alpha+1} = \mathcal{P}(V_\alpha)$, where $\alpha + 1$ is a successor ordinal.

(iii) $V_\lambda = \bigcup_{\alpha<\lambda} V_\lambda$ (if λ is a limit ordinal)

The structure of the V_α thus captures the idea that we take all possible subsets at each additional stage (i.e. iterate powerset) and collect them together at limits (i.e. take a union).

The weak iterative conception is in some ways less well studied than the strong iterative conception, possibly partly because the latter is seen as the default. However, since the weak iterative conception is more general and will be important later, it will be worth getting it on the table (we will provide some more details in Chapter 7).

We'll see a few examples of the weak iterative conception in this book, but some will have to wait until we have a couple of set-theoretic constructions under our belt. For now, here's an easier example to get a feel for it. Suppose we want to build the hereditarily finite sets (i.e. finite sets that are built up out of only finite sets all the way down—formally we say that the empty set is hereditarily finite, and any other set is hereditarily finite just in case it is finite and all its members are hereditarily finite). In standard set theory, we can get these sets just by taking powersets from the empty set (i.e. moving up through each V_n for every natural number n). But there are other ways we might build these sets. Suppose we individuate sets in stages by starting with the empty set at stage 0 and forming at stage $n + 1$ all sets of size at most n. At as we continue through all the stages up to ω (the first infinite stage), we'll eventually get every hereditarily finite set. But we won't get every possible subset at a successor stage. For instance you can check that stage 4 has eight members, so you'll miss out some subsets of stage 4 when moving to stage 5 (you'll have to wait until stage 8 before you can form all subsets of stage 4). So this procedure is weakly but *not* strongly iterative—there are possible sets that don't get formed at the next stage.

We can also have processes that are not even *linearly ordered*, for instance by having two or more set forming operations. For example, let the operation **Even!** form the subsets of a stage with an even number of elements. The other **Odd!** forms the odd numbered subsets of a given stage. By interleaving **Even!** and **Odd!** finitely many times we can get any hereditarily finite set. But the process is not linearly ordered, for instance we could choose to do **Even!** a bunch of times in a row. One doesn't even have a guarantee that you get every hereditarily finite set using these processes (say if you just head off only iterating **Even!** over and over again).

[7]( ) For simplicity, I am giving the version for pure sets, if you want to include Urelemente then clause (ii) should be replaced by $V_{\alpha+1} = \mathcal{P}(V_\alpha) \cup V_\alpha$.

There are more mathematically interesting kinds of weakly iterative process and conceptions of set. Here's a more difficult (but important) example:

Definition 21. (Informal) The *constructibilist conception* holds that sets are formed in stages. Sets are formed by collecting together sets at previous stages that are *definable* (i.e. can be picked out by a formula) over that stage. There are no other sets beyond those obtained this way.

Is conception weakly or strongly iterative? We can show that there are versions of it that are only *weakly* iterative.

( ) Often set theorists will talk about the *constructible universe* (or L) and *constructible hierarchy*. L is formed by taking *definable* powersets. A subset x of the domain of a structure $\mathfrak{M}$ is *definable over* $\mathfrak{M}$ iff x is the unique set containing all and only the y in the domain of $\mathfrak{M}$ satisfying $\phi(y)$ (in $\mathfrak{M}$) for some condition $\phi(y)$ in the language of $\mathfrak{M}$.[a] For a structure $\mathfrak{M}$, let's call the collection of all such $\mathfrak{M}$-definable subsets $Def(\mathfrak{M})$. Then L can be defined as:

Definition 22. The *constructible hierarchy* (or just L) is defined as follows:

(i) $L_0 = \emptyset$

(ii) $L_{\alpha+1} = Def(L_\alpha)$ for successor ordinal $\alpha + 1$

(iii) $L_\lambda = \bigcup_{\alpha<\lambda} L_\alpha$ for limit ordinal λ.

The axiom that every set is constructible i.e. 'For every x there is an α such that $x \in L_\alpha$' is called the *Axiom of Constructibility* or $V = L$.

Now the constructible hierarchy clearly satisfies the weak iterative conception and the constructibilist conception. But it *doesn't* satisfy the strong iterative conception. This is because often new subsets of previous levels get formed as we climb. For example, new subsets of ω coding new real numbers get formed as we move up through the first few stages above L_ω. To see this, note that a *satisfaction predicate* for L_α is definable over $L_{\alpha+1}$, and above V_ω these will code new subsets of natural numbers. This phenomenon (the slow growth of L) is quite general. Since there are only as many formulas as there are parameters available (the usual formula-building operations are trivial at infinite cardinals) we have that the cardinality of L_α is the same as the cardinality of $L_{\alpha+1}$ for every α (in stark contrast to the V_α-hierarchy where $V_{\alpha+1}$ is *always* bigger than V_α). So the L_α hierarchy does not satisfy the strong iterative conception, there are possible subsets that don't get picked up when we move to a successor stage.

Moreover, we could make the iteration more fine-grained and non-linearly-ordered. I could take each *formula* to provide its *own* set-forming operation, and think of successively forming subsets *for specific formulas*, instead of taking the whole definable powerset. This would still qualify as weakly iterative.

Note: Sometimes you can *recover* a version of the strong iterative conception from the weak one. In the case of our n-sized-set-forming operation, we could eventually recover the V_n-hierarchy if we wait long enough. This holds for the L_α-hierarchy too, for example if L satisfies ZFC, it can *recover* its own version of the V_α-hierarchy. However, one still sees the difference between the two hierarchies, even when we assume that V is equal to L and that ZFC holds, it is not the case that $V_\alpha = L_\alpha$ for every α (the L_α-hierarchy takes time to 'catch up').

[a] This is fiddly to formulate. See Chapter 3, §5 of [Drake, 1974].

So there are really multiple sharpenings of the weak iterative conception. One is the strong iterative conception. But others (e.g. the constructibilist conception) are mathematically interesting and not strongly iterative. Clearly, some of these conceptions can be used to motivate certain axioms (for example, the constructiblist conception motivates the Axiom of Constructibility). But what else is there?

4.3 ( ) A modal stage theory for the strong iterative conception

We want our conceptions of set to motivate virtuous theories. Later (Chapter 7) we'll see how versions of the weak iterative conception other than the constructibilist conception can be used to do just this. For now, we'll focus on the 'default' strong iterative conception and ZFC. In particular, I'll:

(1.) Explain ZFC set theory.

(2.) Show how ZFC can be motivated on the basis of a modal axiomatisation of the strong iterative conception.

So, let's start by setting up ZFC:

Definition 23. *Zermelo-Fraenkel Set Theory with the Axiom of Choice* (ZFC) is formulated in the language of set theory $\mathscr{L}_\in$. It comprises the following axioms (we just give informal statements, formal definitions are available in many set theory textbooks):

(i) *Axiom of Extensionality.* Sets with the same members are identical.

(ii) *Axiom of Pairing.* For any two sets x and y there is a set containing just x and y.

(iii) *Axiom of Union.* For any set x, there is a set of all elements of members of x.

(iv) *Powerset Axiom.* For any set x, there is a set of all subsets of x.

(v) *Axiom of Foundation.* Every set contains an element that is disjoint from it. The axiom both rules out self-membered sets and also the existence of infinite descending membership chains.

(vi) *Axiom of Infinity.* There's a non-empty set x such that for any member y of x there is another member z of x such that y is a member of z. (This guarantees that there's an infinite set.)

(vii) *Axiom Scheme of Replacement.* If a formula $\phi(x, y)$ is function-like (i.e. for any x, there is exactly one y such that $\phi(x, y)$, then the image of any particular set under $\phi(x, y)$ is also a set.

(viii) *Axiom Scheme of Separation.* If $\phi(x)$ is a formula in one free variable x, then if y is a set, then there's a set of all the x in y such that $\phi(x)$ (i.e. $\{z|z \in y \wedge \phi(z)\}$ exists).

(ix) *Axiom of Choice.* (AC) For any non-empty set of pairwise-disjoint non-empty sets, there is a set that picks one member from each. (**Note:** ZFC without AC is just denoted by "ZF".)

As noted earlier (Chapter 2) ZFC is a very nice theory of sets with many theoretical virtues. But can it be motivated using the iterative conception?

There's different ways to do this. One way is to axiomatise the notion of a stage *directly*.[8] A different way (and the approach we'll focus on here) is to think of the iterative conception as describing a particular kind of *modal* framework where stages are worlds, and whatever set-forming operations we have provide accessibility. This is a relatively old idea, going back to [Parsons, 1983], but the idea has been fruitfully applied recently. In particular, [Linnebo, 2013] shows how one can give a modal version of the strong iterative conception that motivates ZFC. Giving the full details would take up too much space, but a flavour of the approach will be useful.[9]

We'll want to talk about reifying pluralities into sets, and for this Linnebo uses a plural logic. Really though, any extensional second-order variables would do. Since much of the literature (e.g. [Scambler, 2021]) follows this convention

[8] See here [Button, 2021a] for a recent article on the state of the art.

[9] Details can be found in [Linnebo, 2013] and [Scambler, 2021], and a different modal approaches in [Studd, 2013] and [Button, 2021b].

of using plurals, we'll stick with it. Again we'll leave the plural logic relatively informal, the reader wishing to see a concise presentation of the details is directed to [Linnebo, 2014] or [Oliver and Smiley, 2013] for a textbook treatment. Plural logic has new variables xx that range over 'some things' (e.g. the books on my table), a binary relation symbol $\prec$ (where $x \prec xx$ is to be read as "x is one of the xx"), with the expected definition of well-formed formula. We'll denote the language obtained by adding these resources to $\mathscr{L}_{\in}$ by "$\mathscr{L}_{\in,\prec}$". We'll routinely abuse singularisation and speak of "a plurality" (a standard move in this field).

For our plural axioms (here we're mostly following the presentation in [Scambler, 2021]) we'll take the following:

Definition 24. *Extensional plural logic* has the axioms (again, we give axioms informally, suppressing the formal details, see [Linnebo, 2014]):

(i) A principle of extensionality for plurals (that if two pluralities xx and yy comprise the same things, then anything that holds of the xx also holds of the yy and vice versa).

(ii) An impredicative comprehension scheme:

$$\exists xx \forall y \big(y \prec xx \leftrightarrow \phi(y)\big)$$

for any ϕ in $\mathscr{L}_{\in,\prec}$ not containing xx free.

We then need a background modal logic to talk about moving between the stages. For this we'll add a modal operator $\Diamond$ to $\mathscr{L}_{\in,\prec}$ to get a language $\mathscr{L}^{\Diamond}_{\in,\prec}$, with well-formed formulas as normal. We'll also use the modal operator $\Box$, and in this context $\Box\phi$ can be treated as shorthand for $\neg\Diamond\neg\phi$. For modal axioms we'll use:

Definition 25. Classical $\mathsf{S4}$ is the modal logic with an additional modal operator $\Diamond$ and the axioms:

(i) The necessity of identity and distinctness (these are sometimes optional, but we'll include them):

- $x = y \rightarrow \Box(x = y)$
- $x \neq y \rightarrow \Box(x \neq y)$

(ii) K: $\Box(\phi \rightarrow \psi) \rightarrow (\Box\phi \rightarrow \Box\psi)$ (this holds for any normal modal logic).

(iii) T: $\phi \rightarrow \Diamond\phi$ (this holds if the accessibility relation is reflexive).

(iv) $\mathsf{4}$: $\Diamond\Diamond\phi \rightarrow \Diamond\phi$ (holds if the accessibility relation is transitive).

To obtain $\mathsf{S4.2}$ we add:

(v) G (sometimes called .2): $\Diamond\Box\phi \rightarrow \Box\Diamond\phi$ (holds if the accessibility relation is directed).

The logic S4.3 is obtained by adding:

(vi) .3: $(\Diamond\phi \wedge \Diamond\psi) \rightarrow \Diamond\big((\Diamond\phi \wedge \psi) \vee (\phi \wedge \Diamond\psi)\big)$ (holds if the accessibility relation is linear).

Throughout, we will also assume:

(vii) The Converse Barcan Formula (CBF): $\exists x\Diamond\phi \rightarrow \Diamond\exists x\phi$ (this can be thought of as capturing the idea that domains only grow).

Because we have S4.2 you can think of the space of worlds as a kind of branching time structure, but where you can always bring together any two possibilities (this is the content of the G/.2 axiom). Thus $\Box\phi$ can be though of as saying "in all future worlds ϕ" and $\Diamond\phi$ as "there is a future world such that ϕ".

Before we give the axioms for the stages, we should clarify how we'll interpret *non*-modal set theory. Mostly mathematicians will just want to work with a non-modal axiomatisation of sets, without paying attention to finicky modal details about how the sets are formed in stages. So we can ask: Is there a way of interpreting non-modal set theories in $\mathscr{L}_{\in}$ into our modal language $\mathscr{L}^{\Diamond}_{\in,\prec}$? Given the iterative conception, how should we interpret the 'usual' quantifiers $\forall$ and $\exists$? Well, one natural thought is that $\forall x\phi$ should hold if no matter how you form sets, ϕ will *always* hold, and $\exists\phi$ tells you that you *can* go on to form sets such that ϕ. We can then provide:

Definition 26. Given a sentence ϕ in $\mathscr{L}_{\in}$, the *potentialist translation of* ϕ (denoted "$\phi^{\Diamond}$") is obtained by replacing every universal quantifier "$\forall$" by "$\Box\forall$", and every existential quantifier "$\exists$" by "$\Diamond\exists$".

We can then define a version of the modal stage theory axioms that is extracted from [Linnebo, 2013]:[10]

Definition 27. Lin is the following theory in $\mathscr{L}^{\Diamond}_{\in,\prec}$:

(i) Classical first-order predicate logic.

(ii) Extensional plural logic.

[10] I'm basically following the presentation in [Scambler, 2021], with a few extra tweaks that will be useful later. Scambler uses "L" to denote Lin, I've opted for syntax that avoids possible confusion of Lin with the constructible hierarchy L. Strictly speaking [Linnebo, 2013] doesn't include the plural version of the Axiom of Choice (he is looking for interpretation with ZF, which is all we need for getting the V_{α}s) but [Scambler, 2021] does (but he throws it in as part of the plural logic, I include it as a hybrid plural-cum-set-theoretic axiom). With these systems you get as much Choice out as you're willing to throw in, and since we're primarily interested in ZFC in this book, I'm happy to throw it in.

(iii) Classical $\mathsf{S4.2}$ with the Converse Barcan Formula added.

(iv) The Axiom of Foundation (rendered as normal using solely resources from $\mathscr{L}_\in$).

(v) Extensionality (again using solely resources from $\mathscr{L}_\in$).

(vi) Stability axioms for $\prec$ and $\in$ (these mirror the necessity of identity/distinctness):

- $x \in y \to \Box(x \in y)$
- $x \notin y \to \Box(x \notin y)$
- $x \prec yy \to \Box(x \prec yy)$
- $x \not\prec yy \to \Box(x \not\prec yy)$

(vii) ($\mathsf{Collapse}^\Diamond$) The principle that any things (at a stage) could form a set:

$$\Box\forall xx\Diamond\exists y\Box\forall x(z \in y \leftrightarrow z \prec xx)$$

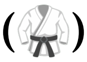
(viii) The axiom that there could be some things comprising all and only the natural numbers.

(ix) The axiom that there could be some things that are all and only the subsets of a given set.

(x) Every potentialist translation of the Replacement Scheme of ZFC.

(xi) A plural version of the Axiom of Choice 'For any pairwise-disjoint non-empty sets xx, there are some things yy that comprise exactly one element from each member of the xx'.

Together, these axiomatise the modal process of the strong iterative conception. Given a world, there could be a plurality of all subsets of that world, and these can then be reified into a set using $\mathsf{Collapse}^\Diamond$. Lin thus gives a way of getting at the strong iterative conception. But how does it relate to ZFC?

4.4 () Mirroring theorems

The core concept will be the idea of *Mirroring Theorems*. These tell you how you can go between the modal theories and the non-modal theories favoured by mathematicians. In particular we can show:

Theorem 28. [Linnebo, 2010], [Linnebo, 2013] ZFC proves ϕ iff Lin proves $\phi^\Diamond$.

This theorem shows that the modal idea of reifying all pluralities into sets at a stage (and continuing this into the transfinite) motivates ZFC concerning the sets. Moreover, it shows how Lin is strongly faithful to 'normal' set theory under the potentialist translation.[11]

Interestingly, the relationship goes back the other way too. Earlier, we mentioned that the strong iterative conception suggests that the universe is formed via the V_α hierarchy. But one can also show:

Theorem 29. [Linnebo, 2013] Over a model of ZFC, the V_α under $\subseteq$ provide a model for Lin.[12]

This shows that not only does Lin motivate ZFC, but if you accept ZFC then you can also get a model for the modal process axiomatised by Lin.

One final piece of the puzzle ties everything together:

Theorem 30. (ZF) For every set x there is an ordinal α such that $x \in V_\alpha$.

This theorem shows you that not only does ZFC allow you to *define* the V_α, but you can prove that *every* set is contained therein. We're now in a position where (i) the strong iterative conception (as axiomatised by Lin) motivates ZFC, (ii) ZFC allows you find a model for Lin, and (iii) ZFC proves that every set lives in said model.[13]

So modal theories of the strong iterative conception and 'standard' set theory, even if it's perhaps too strong to say that they're two sides of the same coin, nonetheless fit *very* well together. Modal stage theory, suitably formulated, pushes the idea that ZFC should be true of the sets, and if ZFC is adopted, we can show that a sensible modal stage theory is a *mathematical fact of life*—if you have ZFC you also have the strong iterative conception and all the sets live there. Finally, the picture explains why **Indefinite Extensibility** holds and **Universality** fails—the universal set never gets formed because at no stage is there a plurality of all possible sets, at the next stage the powerset operation always forms something new.

---

[11] ( ) Indeed, this result can be strengthened to apply to very weak theories, Tim Button (in [Button, 2021a] and [Button, 2021b]) has shown that one can go back and forth between *tremendously* weak (i) theories of sets, (ii) theories of stages, and (iii) modal stage theories (the theories in question do not even imply there *are* any sets!).

[12] ( ) specifically a Kripke frame validating S4.3.

[13] ( ) Of course these models will be proper-class-sized, so not 'models' in the ordinary sense of the term. I suppress these metamathematical details.

Chapter 5

Forcing as a process of construction

In this chapter, I want to outline forcing in set theory, a way of adding *subsets* to models. Thorough presentations are available in a wide variety of mathematical texts and full detail would just bog down the reader in a book of this kind, so my focus is on giving the informal ideas.

There's two main reasons to go into depth on this topic. First, we'll use forcing to articulate the further versions of the weak iterative conception that we'll consider later. Second, forcing is tremendously important for understanding much of the contemporary literature on the philosophy of set theory and the intuitions that underlie much work in this field. So, having a good grasp of it is no bad thing.

5.1 Forcing: The rough idea

A helpful way to understand forcing is by analogy with *field extensions*. Consider the relationship between the fields of real numbers $\mathbb{R}$ and complex numbers $\mathbb{C}$. One way of thinking of obtaining $\mathbb{C}$ from $\mathbb{R}$ is via the idea of *algebraic closure*. Intuitively speaking, we throw in solutions for $\sqrt{-1}$, and then by closing under the field operations, obtain $\mathbb{C}$.

Forcing is very similar. In fact, according to Cohen (a father of the technique), this analogy was part of his discovery.[1] To see this analogy, let's start by considering the problem forcing was developed to solve. In particular, we were trying to prove that the continuum hypothesis is independent from ZFC. Since we knew that given a model M of ZFC, CH is true in the constructible universe of M (a fact proved by [Gödel, 1940]) one way to proceed was to find a way of making a model of $\neg\mathsf{CH}$ from one satisfying CH. (One could then infer by the

[1] See [Cohen, 1963, p. 113] and [Cohen, 2002, pp. 1091, 1093]. Thanks to Carolin Antos for some discussion of the history here.

Completeness Theorem that neither CH nor its negation followed from ZFC, assuming ZFC consistent.) Since we also knew that (again proved in [Gödel, 1940]) L was the smallest inner model (i.e. transitive model containing all ordinals) under inclusion, the natural idea was to break CH by *adding* sets—much like we could find a root for -1 by moving from $\mathbb{R}$ to $\mathbb{C}$. And this is just what Cohen did in [Cohen, 1963].

To think what we need to break CH, it's helpful to think about what CH and $\neg$CH say about sets of reals and functions. CH recall, says that every set of reals (i.e. something with cardinality no bigger than $2^{\aleph_0}$) is either countable or the same size as $2^{\aleph_0}$. In this way, it says that there are lots of *kinds of function* compared with the *kinds of sets of reals*—every infinite set of reals has a function that either bijects it with $\aleph_0$ (the cardinality of $\mathbb{N}$) or $2^{\aleph_0}$ (the cardinality of $\mathbb{R}$).

$\neg$CH by contrast, says that there are lots *kinds of sets of reals* as compared with *kinds of function*—there's some infinite set of reals x for which there's no bijection between x and $\aleph_0$, but also no bijection between x and $2^{\aleph_0}$.

Let's suppose then that we're given a model M of ZFC + CH. What could we do to break CH? Well, we need to (i) *add* some set x to M, whilst (ii) making sure that we preserve the axioms of ZFC when we add x, and (iii) having a set of reals y in the new model such that there's no bijection between either y and the new set of all reals or natural numbers. This what Cohen showed was possible with forcing: Assuming ZFC is consistent, there's a model M satisfying ZFC (by Completeness). Either (i) M satisfies $\neg$CH (in which case we're done) or (ii) M satisfies CH. If (ii), we can then add a bunch of reals G to M, and close under definable operations to form an extension $M[G]$ satisfying ZFC. In this new model $M[G]$, you can show that the *old* set of reals from M is a set of reals that is neither bijected with $\aleph_0$ nor $2^{\aleph_0}$ in $M[G]$.

If you haven't encountered forcing much before, I want the reader to now stop and pause to think about how, given the rough idea of forcing, we might be able to take a model of $\neg$CH and make CH true again by *adding sets*. What kind of set could we add to a model of $\neg$CH in order to restore CH again (and what would we have to simultaneously *avoid* adding)? (Bear in mind that you can't add natural numbers by forcing—a student once made the ingenious suggestion to me that we bump up the size of $\aleph_0$. Alas, this doesn't work since forcing keeps models transitive, and the natural numbers are isomorphic in all transitive models of set theory.)

The answer is that we need to add *functions* that provide the relevant bijections between the old sets of reals and either $\aleph_0$ or $2^{\aleph_0}$, and do so (i) without *adding reals*, and whilst (ii) preserving ZFC. Again, Cohen showed that forcing lets you do this. Given an M satisfying $\neg$CH, one can *collapse* the cardinals between $\aleph_0$ and $2^{\aleph_0}$ to $\aleph_0$ by adding a set H that allows you to get surjections from the naturals numbers to these cardinals. In the new model $M[H]$, CH is true, since there are now bijections between $\aleph_0$ and the old 'cardinals' between $2^{\aleph_0}$

and $\aleph_0$ (i.e. things that were cardinals between $\aleph_0$ and $2^{\aleph_0}$ in the ground model).

These two kinds of forcing are sufficient to show the following:

Theorem 31. Given a model M of ZFC, so long as we can do forcing over M, then M has:

(1.) An extension $M[G]$ such that $M[G]$ satisfies $\neg$CH. This can be done using forcing that *collapses no cardinals*—it does not add new bijections that make any set look smaller than before.

(2.) An extension $M[H]$ such that $M[H]$ satisfies CH. This can be done by forcing whilst *adding no new reals*—we don't add any new subsets of the natural numbers.

In this sense CH is like a set-theoretic light switch as regards forcing—we can flip it on and off at will by successively forcing to add new sets, and all whilst preserving ZFC.[2] Indeed, forcing is *incredibly flexible*. An example that will be important for us is the following:

Theorem 32. Assume that we can always force over M. Then for any set x in M, there is a forcing extension $M[G]$ in which x is countable.

As above, the idea for proving this theorem is just to add a surjection from $\aleph_0$ to x.

Forcing thus provides us with a very controlled way of adding subsets to models. We'll discuss this a little in a -section below (§5.2), but it will be helpful to indicate the shape of what is to come. Forcing, I want to contend, can be thought of as a *process* for adding subsets to a universe and in particular might be a way of generating sets under the *weak* iterative conception. Using this idea, we'll end up with the motivation of a concept of set on which every set is countable, since given a set x at some stage, we could always add a function by forcing that makes x countable.

5.2 ( ) A little more depth on forcing

In this section I add a little more mathematical detail and provide an intuitive characterisation of forcing. This whole section is a -section, so the reader shouldn't get bogged down in the details unless they really want to. Still, the section will help inform the idea that we can think of forcing as a set-forming process, so I recommend at least giving it a go. Good introductions to this material can be found in [Kunen, 1980] and its update [Kunen, 2013] (a wonderful pair of books explaining a range of issues in detail), [Drake and Singh, 1996] (a

[2] This terminology of 'switches' is from [Hamkins and Loewe, 2008].

nice concise introduction), and [Weaver, 2014] (a much easier-going introduction before the applications starting in Ch. 14). Many set theory texts contain an introduction, however, and the reader should feel free to shop around.

We'll start with an example that will help us follow what comes later a bit better. We'll take the idea of *adding a Cohen real*. Let's suppose that you're in a model of ZFC. For now, we'll assume that the model is countable (and transitive) and so (by Cantor's Theorem for the reals) misses out a whole bunch of real numbers. For our purposes, you can think of a real number as an infinite ω-length sequence of 0s and 1s (this, in turn, can be thought of as a function from the natural numbers into $\{0, 1\}$, which says whether there's a 0 or a 1 in the n^{th} place). I want to now add in a new real number, and do so in such a way that ZFC is satisfied. So I slowly go through deciding on what I want in the n^{th} place of my new real for each n (perhaps not in order). I need to do two things (i) make sure I'm avoiding the reals of M (i.e. I don't get something I already have), and (ii) make sure that when I'm done I close under new definable operations to ensure ZFC is true. This is what forcing lets you do. Such an object (a new ω-length sequence of 0s and 1s) is our new real number (our 'Cohen real').

Let's now take a little peek into the machinery of how we do this. The way I suggest thinking of forcing is as a way of talking about descriptions of collections that can change their members as we make certain decisions. In the end, if we make decisions in exactly the right way, we'll end up defining a new object that isn't currently in the universe we start in, and fill in all the needed sets to make ZFC true. The rough ingredients of forcing are the following (i) a *partial order* $\mathbb{P} = (P, <_{\mathbb{P}})$ with certain nice properties that make it sufficiently 'interesting'. You can think of $\mathbb{P}$ as the space of possible 'decisions' that we might take. (ii) $\mathbb{P}$-names, these are descriptions of collections that can change their membership depending on what decisions we take from $\mathbb{P}$, (iii) *dense sets*, these are like *advisors*, no matter what decisions you've taken, they'll always recommend at least one more you might go on to take, and (iv) a *generic filter*, this you can think of as a complete description of all the decisions that were taken in the limit, consistent with every recommendation given by an advisor. Let's look at these in more detail.[3]

First, we need the notion of a *forcing partial order* $(\mathbb{P}, \leq_{\mathbb{P}})$. Before we give the definition, a couple of notes are in order:

- **Note 1:** We often refer to elements of the partial order as 'conditions'.
- **Note 2:** Here the partial order grows 'downwards'—the intuition being that if $p <_{\mathbb{P}} q$, you've got a smaller range of possible decisions after p

[3]**Note:** Often authors (e.g. [Drake and Singh, 1996], [Weaver, 2014]) write in information-theoretic terms, $\mathbb{P}$ is a space of *information*, and we slowly get more and more fine-grained information as we move through $\mathbb{P}$. The way I'm expressing things is essentially equivalent, but a bit easier to think about philosophically, and brings the 'variable set' way of thinking to the fore a little more.

as compared to q. Some people write $p >_{\mathbb{P}} q$ to indicate the same state of affairs, the intuition being that you've got *more* information from q as compared to p.[4]

We now define:

Definition 33. A *forcing partial order* $\mathbb{P} = (P, \leq_{\mathbb{P}})$ is a partial order $\mathbb{P}$ such that:

(i) $\mathbb{P}$ has a maximal condition $1_{\mathbb{P}}$

(ii) $\mathbb{P}$ is *atomless*—any element of p of $\mathbb{P}$ has incompatible extensions (i.e. there's $q \leq_{\mathbb{P}} p$ and $r \leq_{\mathbb{P}} p$ such that there's no s with $s \leq_{\mathbb{P}} q$ and $s \leq_{\mathbb{P}} r$).

The way I'm going to suggest one thinks about this partial order is as an information space of *possible decisions* for settling membership facts. As we'll see, we can define a class of 'names' for possible sets (these are called $\mathbb{P}$-names). These we can think of as having their membership facts settled as we take decisions through $\mathbb{P}$. The conditions of being atomless one can think of as a condition on $\mathbb{P}$ being sufficiently *interesting* or *non-trivial*—there's always incompatible decisions one could make about where to go, and there's no part of $\mathbb{P}$ that admits of 'inevitability'.

In the specific case of adding a Cohen real, we can define the following partial order:

Definition 34. Given some model M, the forcing partial order to *add a Cohen real* has as its domain (in M) all partial functions from ω into $\{0, 1\}$ and $p \in P$ extends q (i.e. $p \leq_{\mathbb{P}} q$) iff p extends q as a function (i.e. q's domain is a proper subset of p's, and they agree on all arguments from q's domain).

This order gives us a way of thinking of settling the n^{th} place of a new real—as we move down through $\mathbb{P}$ we settle more and more values for a new real to be defined. In the limit, we'll have settled every bit of the real.

How to get a handle on this idea of 'settling values'? For this we'll need the definition of a $\mathbb{P}$-name. The definition looks somewhat complicated, but it can be given an intuitive backing.

Definition 35. A $\mathbb{P}$-*name* is a relation τ such that $\forall \langle \sigma, p \rangle \in \tau($'$\sigma$ is a $\mathbb{P}$-name' $\wedge p \in P)$. In other words, τ is a relation that relates $\mathbb{P}$-names to conditions of $\mathbb{P}$.

The definition *looks* circular, but in fact is not since the empty set is trivially a $\mathbb{P}$-name. You can think of the $\mathbb{P}$-names as relations where other $\mathbb{P}$-names are related to conditions in $\mathbb{P}$.

[4] See [Drake and Singh, 1996], p. 155, Warning 8.8.2 for discussion.

The intuition to have in mind is that a $\mathbb{P}$-name is the name for a possible set. Given a bunch of good 'decisions' from $\mathbb{P}$ (we'll talk about this idea of 'a bunch of good decisions' in a second, the key notion is that of a *generic filter*) we'll evaluate the $\mathbb{P}$-names to different sets in the extension. The way this works is given a $\mathbb{P}$-name σ, we're going to rule in or out other $\mathbb{P}$-names in domain of σ according to whether or not they're related to a condition in our new object (these names will in turn have been evaluated according to different decisions). So $\mathbb{P}$-names are kind of 'variable collections'—they can change their mind as to what they contain as we move about in $\mathbb{P}$.[5]

The next notion we need is:

Definition 36. We say that $D \subseteq \mathbb{P}$ is *dense* iff for every $p \in \mathbb{P}$, there is a $q \in D$ such that $q \leq_{\mathbb{P}} p$.

The way of thinking about a dense set D is that it's kind of like a *advisor*. No matter where you are in $\mathbb{P}$, and what decisions you've taken, D can come up with at least one decision you could take to continue.

Next we need the notion of a *generic filter*:

Definition 37. $G \subseteq \mathbb{P}$ is a *filter on* $\mathbb{P}$ iff:

(i) G is non-empty.

(ii) $p \in G$ and $q \geq_{\mathbb{P}} p$ implies that $q \in G$ (i.e. G is closed upwards).

(iii) $p \in G$ and $q \in G$ implies that there is an $r \leq_{\mathbb{P}} p, q$ with $r \in G$ (i.e. G brings any two elements together).

We furthermore say that G is M-$\mathbb{P}$-generic (for some model M) iff G intersects every dense set of $\mathbb{P}$ in M. (We'll often just abbreviate this to 'generic' and let context determine the values of $\mathbb{P}$ and M.)

The way to think of such a G is that it is a kind of 'maximal' collection of 'good decisions made'. If you include a decision $p \in G$, then you've got to include any earlier decisions that could have lead there, and also you've also got to bring together any two decisions together later (there's no including incompatible decisions allowed). You've also got to be 'good' in that you agree with every advisor (i.e. dense set) in at least one place. Part of what genericity ensures is that you don't encode any 'extra' information in what you add.

We can then talk about what happens to a $\mathbb{P}$-name when presented with a generic G.

Definition 38. We evaluate $\mathbb{P}$-names by letting the value of τ under G (written '$val(\tau, G)$' or 'τ_G') be $\{val(\sigma, G) | \exists p \in G(\langle \sigma, p \rangle \in \tau)\}$.

[5]( ) Interestingly the idea of 'variable collection' corresponds well to the category-theoretic approach to forcing. See the Appendix to [Bell, 2011].

Again, this looks complicated, but the intuition is as follows. Remember that a $\mathbb{P}$-name can be thought of as a kind of 'variable collection' or 'name for a possible set'. When we give some G to a $\mathbb{P}$-name τ, we evaluate stepwise by analysing the valuation of all the names in the domain of τ and then we add them to τ_G according to whether they're related to some $p \in G$. In particular, if σ is a $\mathbb{P}$-name in the domain of τ, then we put σ_G into τ_G if there is a $\langle \sigma, p \rangle \in \tau$ for which $p \in G$ (and throw σ_G out of τ_G otherwise). So you can think of us running through the $p \in G$ and throwing in or out already evaluated $\mathbb{P}$-names according to whether a name is related to some $p \in G$.

Let's return to our example of adding a Cohen real. Consider the following conditions from the poset to add a Cohen real:

- f is defined by:
  - $f(0) = 1$
  - $f(3) = 0$
- g is defined by:
  - $g(0) = 0$
  - $g(3) = 0$

Now consider the following names:

- $\tau = \emptyset$
- $\sigma = \{\langle \tau, f \rangle\}$
- $\mu = \{\langle \tau, f \rangle, \langle \sigma, g \rangle\}$
- $\nu = \{\langle \tau, f \rangle, \langle \tau, g \rangle, \langle \sigma, f \rangle, \langle \sigma, g \rangle, \langle \mu, f \rangle, \langle \mu, g \rangle\}$

Let's suppose that $f \in G$ but $g \notin G$. So this says that the first bit of our new real is 1, and the third bit is 0. What happens to our $\mathbb{P}$-names under G? Well, τ is trivial and so remains unchanged. We now have a value τ_G for τ, so the values σ, μ, and ν will contain $\tau_G = \emptyset$ (since we have $\langle \tau, f \rangle \in \sigma, \mu, \nu$). The evaluation of σ is now complete, and we know that $\sigma_G = \{\emptyset\}$. For μ, since we know $g \notin G$, we *throw out* the evaluation of σ from μ_G, and so $\mu_G = \{\tau_G\} = \{\emptyset\}$. For ν, whilst we do have a bunch of $\mathbb{P}$-names correlated with g (and so the evaluation of those names don't make it in via any ordered pair of the form $\langle \xi, g \rangle$) we also have that ν contains $\langle \tau, f \rangle$, $\langle \sigma, f \rangle$, and $\langle \mu, f \rangle$ and so the interpretation of these names gets thrown in. So $\nu_G = \{\tau_G, \sigma_G, \mu_G\} = \{\emptyset, \{\emptyset\}\}$.

Of course things are much more complicated when we move to names with more structure (in particular once you have big infinite names things are going to get more subtle). But I hope the rough idea is clear. We have a 'space of possible decisions' (the partial order $\mathbb{P}$), a bunch of names that can change their mind about what they contain when presented with some 'decisions' from $\mathbb{P}$ (i.e. the $\mathbb{P}$-names), and a bunch of 'advisors' (the dense sets) each of which can always present to you a way of continuing after some point in $\mathbb{P}$. We're then given a 'maximal good bunch of decisions' (the generic G), that agrees with every dense set at some point and lets you find your way through $\mathbb{P}$ by giving you conditions from $\mathbb{P}$. G tells each $\mathbb{P}$-name who they are by ruling in and throwing out (evaluations of) $\mathbb{P}$-names based on whether the names in the domain of a $\mathbb{P}$-name are related to the decisions in G.

Other partial orders that are especially important are:

Definition 39. Forcing to add κ-many Cohen reals.

- P is the collection of all finite partial functions (in M) from $\kappa \times \omega$ to $\{0, 1\}$
- $p \leq_{\mathbb{P}} q$ iff p extends q as a function.

A generic for this partial order doesn't just add *a* Cohen real and then close under definability, it adds *κ-many*. One can then show that you don't destroy any cardinals (this is a non-trivial lemma[6]) by adding a generic for $\mathbb{P}$. This then let's you infer (picking big enough κ) that $\neg\mathsf{CH}$ holds in $M[G]$, even if M satisfies CH, all the cardinals between ω and κ in M are now cardinalities between ω and κ of different sets of reals in $M[G]$.

As mentioned earlier, any cardinal can be collapsed to the countable using forcing. This is done using:

Definition 40. The forcing to *collapse κ to ω* is defined by:

- P is the collection of finite partial function from ω into κ.
- $p \leq_{\mathbb{P}} q$ iff p extends q as a function.

A generic for this partial order allow us to get a surjection from ω to κ, and collapse the cardinality of κ (and any sets bijective with κ) to ω.

These represent just a taste of some of the possibilities available using forcing. As Joel-David Hamkins writes (about model-building methods including forcing):

> Set theorists build models to order. [Hamkins, 2012, p. 417]

So forcing is a flexible tool that gives us a way of adding sets to models. There's two points we should note. First:

[6] See, for example, [Weaver, 2014, p. 50], Theorem 13.3.

Fact 41. If $\mathbb{P}$ is a forcing partial order in a model M of ZFC, and G is $\mathbb{P}$-M-generic for $\mathbb{P}$, then $G \notin M$. In particular $\mathbb{P} - G = \{p | p \in \mathbb{P} \wedge p \notin G\}$ is dense (and clearly missed by G).

This fact will be a little important later when we relate 'paradoxes' related to forcing and the Russell/Cantor reasoning (I relegate a proof to a footnote[7]).

Fact 42. Let M be a transitive model satisfying ZFC and let $M[G]$ be the model obtained by evaluating all the $\mathbb{P}$-names for a forcing partial order $\mathbb{P}$ and M-$\mathbb{P}$-generic G. Then $M[G]$ also satisfies ZFC, and in particular $M[G]$ is the smallest transitive model of ZFC containing both every element of M and G.[8]

The strategy for proving this is to 'cook up' $\mathbb{P}$-names that you know (by the genericity of G) will ensure that ZFC is satisfied. But the fact that you get the *smallest* possible extension is is important: It shows that you can think of the addition of a forcing generic G and evaluating the $\mathbb{P}$-names as throwing in G and closing under definable operations—i.e. you don't get any 'extra' sets than what is required to get ZFC by throwing in G so long as G is generic. In this way, the $\mathbb{P}$-names and evaluation procedure conspire to make sure the construction of $M[G]$ is very tightly controlled. This further reinforces the similarity between forcing and more mathematically familiar constructions like obtaining the field of complex numbers from the field of real numbers. There, we take $\mathbb{R}$, throw in i, and close under the usual field operations to get $\mathbb{C}$. Indeed, $\mathbb{C}$ is the *smallest* such field. So with forcing, $M[G]$ is the *smallest* model of ZFC you get by throwing in G and closing under every operation you can define.

Moreover, there's a sense in which finding such a G can be thought of as a *process* in its own way. If we're given a forcing partial order $\mathbb{P}$ and a family of dense sets $\mathcal{D}$ (let's let each dense set D_i in $\mathcal{D}$ be indexed by some i in an index set I), we can think of successively hitting each D_i in such a way that we extend our previous choices. What we obtain in the limit will be a generic that hits every D_i in $\mathcal{D}$.

I hope that the reader finds the above helpful, and in particular it can serve as an intuitive road-map if you want to learn forcing in detail (alongside an introductory text). Before we move on, I want to identify:

> **Main Philosophical Upshot.** You can think of moving from M to $M[G]$ by forcing as a way of *generating new sets*.

[7] *Proof.* Suppose $p \in \mathbb{P}$. We must show that there is $q \in \mathbb{P} - G$ such that $q \leq_{\mathbb{P}} p$. The only non-trivial case is where $p \in G$. Because $\mathbb{P}$ is non-atomic, there are incompatible r and s extending p. But then one of r and s isn't in G—all elements of G are compatible with one another.

[8] See [Kunen, 2013], Lemma IV.2.19.

Chapter 6

A 'new' kind of paradox?

In this chapter I want to argue that there's a tension at the heart at the heart of set theory. We'll then (Chapter 7), explain how this can be resolved into different conceptions, much as we saw with the naive conception of set.

6.1 The forcing-saturated strong iterative conception of set

A popular thought in set theory is some idea of *maximality*, the idea that there should be as many sets as possible.[1] Given the thought that we want the universe to be closed under lots of different kinds of operation, we might think it's natural to hold the following conception of what the sets are like:

Definition 43. (Informal) The *forcing-saturated strong iterative conception of set* holds that sets are formed in stages. There are two operations. One can either (i) form the set of all possible subsets of the stage, or (ii) add in a generic for a partial order and a family of dense sets.

So, we have two operations. We can either add in a forcing generic, or we can form the powerset of a set. Clearly then, we have:

Powerset. The Powerset Axiom holds.

We're going to shortly see some conflicts with **Powerset**. We therefore define:

Definition 44. ZFC with the Powerset Axiom removed and with Separation and Collection (the principle that the range of a a function on a set is *contained* in some set) will be called ZFC^-.

[1] See [Incurvati, 2017] for a survey of this idea in set theory.

( ) **Note:** Dropping the Powerset Axiom is a slightly subtle business. Simply deleting it results in a theory weaker than one would like. Instead, one should substitute schemes of Separation and Collection (the principle that the range of a function is *contained* in some set) for Replacement, as these are not equivalent without Powerset. See [Zarach, 1996] and [Gitman et al., 2016] for discussion.

Since, we can also introduce a generic for any forcing partial order and family of dense sets under the forcing-saturated strong iterative conception, we'll introduce the following axiom:

Definition 45. (ZFC^-) By the *Forcing Saturation Axiom* or FSA we mean the claim that for *any* partial order and set $\mathcal{D}$ consisting solely of dense sets for $\mathbb{P}$, there is a generic G intersecting every member of $\mathcal{D}$.[2]

The forcing saturated strong iterative conception thus motivates:

Forcing Saturation. The Forcing Saturation Axiom holds.

Readers familiar with forcing may already see the problem with the forcing-saturated strong iterative conception. For the reader that isn't, I want them to briefly pause and think about what **Powerset** entails (especially in light of Cantor's Theorem) and what follows from **Forcing Saturation** (especially given collapse forcings).

6.2 The Cohen-Scott Paradox

Here's the problem: The forcing-saturated strong iterative conception motivates both **Powerset** and **Forcing Saturation**, but they're inconsistent with one another. This mirrors how the naive conception was brought down by **Universality** and **Indefinite Extensibility**. I'll refer to the paradox I'll give as the *Cohen-Scott Paradox* as it originates with the mathematical work of Cohen, and Scott was one of the first to propose the tension I'll identify. The paradox is thus not really that 'new', and the idea that there might be a tension between wanting uncountable sets and forcing has been around since at least the 1970s. However, recent work has developed the philosophy and mathematics of these ideas substantially, including [Meadows, 2015], [Scambler, 2021], [Builes and Wilson, 2022], and [Barton and Friedman, MS].[3]

Letting "Powerset" denote the Powerset Axiom, the current theory motivated by the forcing-saturated strong iterative conception is $\mathsf{ZFC}^- + \mathsf{Powerset}$

[2] See [Barton and Friedman, MS], Definition 9.

[3] Naming the problem "The Cohen-Scott Paradox" is taken from [Barton and Friedman, MS].

+ FSA. But we can now note that because we can produce a generic for any forcing partial order and family of dense sets, we can use the collapse forcing to add a generic making any set countable. In fact, we can note the following:

Fact 46. (ZFC^-) The forcing saturation axiom is equivalent (modulo ZFC^-) to the axiom "Every set is countable".[4]

We can now present the Cohen-Scott Paradox:

> **The Cohen-Scott Paradox.** Simply put, ZFC^- + Powerset + FSA implies that there are uncountable sets (by Cantor's Theorem and the Powerset Axiom) but also that every set is countable (by the Forcing Saturation Axiom). Contradiction!

Before we continue, I want to emphasise: No reasonable classical set theorist has ever accepted both **Forcing Saturation** and **Powerset** in this generality. Perhaps someone learning forcing might unwittingly fall into the trap of accepting the forcing-saturated strong iterative conception, or perhaps its appealing to a theorists of a dialethic persuasion. But set theorists are a clever bunch, and they are able to see this contradiction coming a mile off. In fact, this tension has been noticed for a while. Discussing forcing in the introduction to Bell's book on the subject, Dana Scott writes:

> I see that there are any number of contradictory set theories, all extending the Zermelo-Fraenkel axioms: but the models are all just models of the first-order axioms, and first-order logic is weak. I still feel that it ought to be possible to have strong axioms, which would generate these types of models as submodels of the universe, but where the universe can be thought of as something absolute. Perhaps we would be pushed in the end to say that all sets are countable (and that the continuum is not even a set) when at last all cardinals are absolutely destroyed. [Scott, 1977, p. xv]

So the Cohen-Scott 'Paradox' is certainly not new, and was noticed from the inception of forcing. So if it's so obviously bad, why even consider the forcing-saturated strong iterative conception? The reason to do so is not that individual agents hold it, but that it forces us to face a possible *choice*. Much as we saw with the naive conception, there's different ways we could go. We could adopt a version of the logical conception that validates **Universality**. Or we could adopt a version of the combinatorial iterative conception that on which **Indefinite Extensibility** holds. Similarly, we could now adopt **Powerset** (for example by

[4]This is a well-known folklore result, but see Fact 10 of [Barton and Friedman, MS] for details.

holding the strong iterative conception) or we could adopt a conception that validates **Forcing Saturation**. We'll explore this in more detail shortly (in Chapter 7). For now I want to consider the relationship between the Cohen-Scott Paradox and diagonalisation, before we go on to consider how we might resolve the paradox.

6.3 ( ) The Cohen-Scott Paradox and diagonalisation

To see the link with 'diagonal' arguments, we start with the question:

Question. What (if any) is the link between the Russell-Cantor reasoning and the Cohen-Scott Paradox?

We have already seen a tight link between Russell's Paradox and Cantor's Paradox in Chapter 3—in the case where we first take the universal set, then consider the identity surjection/injection, and then run the standard proof of Cantor's Theorem, we get the Russell set.

There is a superficial similarity here, in that the (un)countability of some set x can be viewed as a claim about the (non-)existence of a surjection from ω to x. But is there any deeper similarity?

As mentioned earlier, the assumption that every set is countable (i.e. for any set x there is a surjection from ω to x) is *equivalent* (over ZFC^-) to the claim that for any forcing partial order and any set-sized family of dense sets $\mathcal{D}$, there is a generic intersecting $\mathcal{D}$ (i.e. the Forcing Saturation Axiom).[5]. We can now present the following 'diagonal' version of the Cohen-Scott Paradox.

The Cohen-Scott Paradox, Diagonal Version. If the Powerset Axiom is true, then the family $\mathcal{D}^*$ of *all* dense sets for $\mathbb{P}$ is a set-sized family. By the Forcing Saturation Axiom, there is a generic G intersecting every member of $\mathcal{D}^*$. Now consider $E = \{p | p \notin G\}$. It's an exercise to show that E is dense, the interested reader can go back and find the proof in Chapter 5. Since G is generic for $\mathcal{D}^*$ and E is dense, we know that G intersects E at some point p. But then we have $p \in G \leftrightarrow p \in E$ (by choice of p), but $p \in E \leftrightarrow p \notin G$ (by the definition of E), and so $p \in G \leftrightarrow p \notin G$ (putting together the biconditionals), contradiction![a]

[a] See here also [Meadows, 2015] for emphasis of this diagonal version of the Cohen-Scott Paradox

The point to note here is that there is a similarity to the Russell-Cantor reasoning. There we had the assumption of the existence of a particular surjection

[5] ( ) See [Barton and Friedman, MS], Fact 16.

leading to contradictory claims about (non-)self-membership. Here we have the existence of an surjection, whilst not leading to contradictory claims about *self*-membership, we do have the contradictory $p \in G \leftrightarrow p \notin G$. So whilst the analogy is not perfect, we have a diagonal-style contradiction obtained by assuming the existence of a particular surjection. We'll discuss a possible significance of this in §9.3.

To sum up, we've seen that:

(1.) There is a tension between **Forcing Saturation** and **Powerset**

(2.) This can be put in terms of a diagonal argument, with similarities to the Russell-Cantor reasoning.

So, what to do about *this* state of affairs?

Chapter 7

Countabilist conceptions of iterative set

We've identified a tension between **Forcing Saturation** and **Powerset**, in analogy with **Universality** and **Indefinite Extensibility**. And just as before, we can move forward by dropping one of the two. One way is to just hold that **Forcing Saturation** should be dropped and **Powerset** accepted. This pushes us towards the strong iterative conception and the modal stage theory given by Lin. But might there be a way of going forward with **Forcing Saturation** instead of **Powerset**? In this chapter we'll see some stage theories that validate **Forcing Saturation**. In the next (Chapter §8) we'll discuss how these conceptions interpret mathematics, and compare the two approaches in light of the theoretical virtues adumbrated in Chapter 2.

7.1 Countabilist stage theories

As we'll see shortly, if you're going to have **Forcing Saturation**, then every set is going to be countable. For the sake of brevity, it will be helpful to introduce some terminology:

Definition 47. The *countabilist axiom* (or Count) is the axiom 'Every set is countable'.

Definition 48. (Informal) We will refer to the view that holds Count as *countabilism* (with *countabilist* the corresponding adjective).

It's fair to say that countabilist options for the (weak) iterative conception been a lot less studied than the 'standard' strong iterative conception, and so we will have to proceed with a little more care in articulating the alternative. This way of viewing the sets is still somewhat nascent with much work still to be done, and we will have to be cautious in our conclusions. It is less solidified than the standard strong iterative conception, and I don't want to overstate my case. I *do* want to identify, however, that it's an *attractive* alternative. This section, then, will have the flavour of explaining a *promising road of inquiry*, rather than the highly solidified picture of the strong iterative conception.

Since we have Count for the countabilist, we can't have uncountable sets. For this reason, we're going to have to drop the Powerset Axiom and adopt ZFC^- + Count. Since we don't have the Powerset Axiom (indeed we have its negation) we don't have the V_α, and so we're going to have to adopt the weak iterative conception, rather than the strong iterative conception. So the question then becomes: Given that the V_α are out, what could our stages be, and how are they given? Recall that for any weak iterative conception we need:

(i) A description of what counts as a starting domain.

(ii) A description of some operation(s) for forming new sets from old.

Can we come up with weakly iterative stage theories for the countabilist, and thereby give a story along the lines of (i) and (ii)?

7.2 Reify! and Generify!

I want to argue that there are proposals in the literature that *can* be viewed as providing stage theories for countabilist versions of the weak iterative conception. Before I mention some concrete proposals, I want to identify some motivations in the literature.

Regarding (i): What might the *processes* be? Well, one possibility is familiar—given some stage we want a notion of forming sets out of the classes of that stage. This is what the Powerset Axiom codifies—every possible class at some stage V_α is reified into a set (if it didn't already exist) at $V_{\alpha+1}$ and can be formalised modally by Lin. But note, we don't have to turn *every possible class* in a set at a subsequent stage. This is made clear by the constructibilist conception and the constructible hierarchy, at $L_{\alpha+1}$ we reify those classes *definable* over L_α into sets. For example, we'll get the 'universal class' of the previous stage at the next one, since $x = x$ is a perfectly good formula. So, one class of operations are given by what I'll call ***Reify!*** commands; they take some 'proper classes' of the domain and reify them into sets.

However, as I hope I convinced the reader in Chapter 5, another kind of operation for adding sets is *forcing*. We can thus think of having, in addition to whatever **Reify!** commands we employ, an operation ***Generify!*** which will take in a partial order $\mathbb{P}$ and family $\mathcal{D}$ of dense sets, and spit out a generic for $\mathbb{P}$ and $\mathcal{D}$. Closely linked is the operation ***Enumerate!*** that adds an enumeration between a set and the natural numbers. There are a class of **Enumerate!** commands that can be thought of as special cases of the **Generify!** operation, in particular the specific case of the forcing that adds a surjection from the natural numbers to a set. If we think that the stages should support **Generify!**, then **Enumerate!** will always be executable. This idea has been advocated recently by a few authors. For example, Chris Scambler writes:

> The guiding idea...is to introduce another way of extending a given universe of sets as an option at each stage of the process. Specifically, we will imagine we are capable not only of introducing sets whose members are among already given things..., but also of introducing new functions between already given (infinite) sets, and in particular of introducing functions defined on the natural numbers and whose range contains any set as a subset [Scambler, 2021, p. 1088]

Jessica Wilson and David Builes express a similar idea (partly drawing on [Scambler, 2021]):

> Recall that any set-theoretic universe is ultimately generated by two sorts of processes: the powerset operation and the length of the ordinals. Proponents of height potentialism maintain that the length of the ordinals is indefinitely extensible: necessarily, for any ordinals, there could always be more. The modal approach to [Cantor's Theorem] simply extends this line of thought to the powerset operation: necessarily, for any subsets of an infinite set, there could always be more. This is width potentialism. For any set-theoretic structure, there is both a taller one and a wider one. [Builes and Wilson, 2022, p. 2212]

Recall how we could use **Even!** and **Odd!** to obtain the hereditarily finite sets. Can we think of interleaving **Reify!** and **Generify!** to obtain a stage theory for countabilist set theories? The answer is yes.

7.3 () A reifying and generifying modal stage theory

[Scambler, 2021] has provided a theory of worlds that can be thought of as providing a modal stage theory for countabilist versions of the weak iterative conception. He starts with the background of $\mathscr{L}^{\Diamond}_{\prec,\in}$ but adds two modal operators $\langle v\rangle$ (for 'vertical' modality—reifying the pluralities of the model into sets) and $\langle h\rangle$ (for 'horizontal' modality—adding in subsets via forcing). Call this language $\mathscr{L}^{\Diamond,\langle h\rangle,\langle v\rangle}_{\in,\prec}$. Boxes $[h]\phi$ and $[v]\phi$ are defined as $\neg\langle h\rangle\neg\phi$ and $\neg\langle v\rangle\neg\phi$ as usual. In this context, the general $\Diamond$ can be thought of as 'possible by iterating either operation'.

Scambler then provides the following axioms [Scambler, 2021, p. 1091]:

Definition 49. Sca consists of the following axioms in $\mathscr{L}^{\Diamond,\langle h\rangle,\langle v\rangle}_{\in,\prec}$ (again, I focus on giving more intuitive statements, the reader should go to [Scambler, 2021] for the formal details):[1]

(i) Classical first-order logic.

(ii) Extensional plural logic.

(iii) Classical S4.2 with the Converse Barcan Formula for every modality.

(iv) The necessity of distinctness and stability axioms for $\prec$ and $\in$ (Scambler calls these 'definiteness axioms', but we'll follow [Linnebo, 2013]'s terminology).

(v) The Axiom of Foundation (the standard one from ZFC).

(vi) Extensionality for sets (again, no different from ZFC).

(vii) **Weakening Schemas**: $\langle h\rangle\phi \to \Diamond\phi$ and $\langle v\rangle\phi \to \Diamond\phi$, for every ϕ.

(viii) **Vertical collapse**: $\langle v\rangle\exists y\Box\forall z(z \in y \leftrightarrow z \prec xx)$.

(ix) The axiom that there could *vertically* be some things that necessarily comprise all and only the natural numbers: $\langle v\rangle\exists xx\Box\forall y(y \prec xx \leftrightarrow$ 'y is a natural number').

(x) **Subset Comprehension.** The axiom that its *vertically* possible to have some things that are *vertically necessarily* all the subsets of a set: $\forall z\langle v\rangle\exists xx[v]\forall y(y \prec xx \leftrightarrow y \subseteq z)$.

(xi) **Possible Generics.** The axiom 'If $\mathbb{P}$ is a forcing partial order and dd is some dense sets of $\mathbb{P}$, then it's horizontally possible that there is a filter meeting each dense set that is one of the dd'.

(xii) The plural version of the Axiom of Choice 'For any pairwise-disjoint non-empty sets xx, there are some things yy that comprise exactly one element from each member of the xx' (Scambler throws this in with the plural logic, but as before we'll keep it separate).

[1] Scambler uses the term "M" (for **M**eadows) to denote Sca, as he takes inspiration for his view from [Meadows, 2015]. As we'll see below, Meadows' work (drawing on [Steel, 2014]) is slightly different (he does not have a vertical modality), therefore I've chosen the term "Sca".

Some of these axioms deserve a mention. The **Weakening Schemas** are meant to capture the idea that there if I could get a set by either reifying pluralities into sets or forcing, then such a set is possible simpliciter. **Vertical Collapse** axiomatises the idea, as with Lin, that I could reify any plurality over a world into a set. **Possible Generics** corresponds to the idea that I could always add a generic for any partial order. One issue then is **Subset Comprehension**: Notice that it is restricted to the vertical modality—this will not hold in general since one can always add subsets along the horizontal modality.

Two theorems are especially important for assessing the import and intuitions behind Sca. First we have:

Theorem 50. Sca interprets $\mathsf{ZFC}^- + \mathsf{Count}$ under the potentialist translation (the potentialist translation, recall, takes a formula ϕ in $\mathscr{L}_\in$ to a corresponding one in $\mathscr{L}_{\in,\prec}^{\Diamond,\langle h\rangle,\langle v\rangle}$ by replacing every occurrence of $\forall$ with $\Box\forall$ and every occurrence of $\exists$ with $\Diamond\exists$).[2]

So there's a sense in which when we have the full modality, thinking of the stages as given by Sca gets us $\mathsf{ZFC}^- + \mathsf{Count}$. However we also have:

Theorem 51. [Scambler, 2021] Sca interprets ZFC when we restrict to the vertical modality (i.e. when we do the potentialist translation but replace $\Box$ and $\Diamond$ by $[v]$ and $\langle v\rangle$ throughout).

So when we restrict to the vertical modality under Scambler's stage theory we get ZFC in the non-modal theory (this is basically just because the vertical modality obeys Lin). However, we have to *ignore* the horizontal modality that would allow us to collapse any given uncountable set (and hence break the Powerset Axiom in the non-modal theory).

The intuition behind Sca is thus the following. We have the vertical modality that will allow us, starting with the empty set, to obtain ZFC by successively reifying classes of worlds. However, we could, at any point, choose to introduce a generic for a given partial order and family of dense sets. And, by interleaving **Reify!** and **Generify!** we can get $\mathsf{ZFC}^- + \mathsf{Count}$. Note that, unlike the strong iterative conception or the constructibilist conception, the stages provided by Sca need not be well-ordered. Instead, much like **Odd!** and **Even!**, we have to think of applying **Generify!** and **Reify!** appropriately. Just as if you spin out applying one of **Odd!** or **Even!** you won't get all the hereditarily finite sets, so with **Generify!** and **Reify!**. If you head off applying **Reify!** over and over again, you'll just get ZFC. And there are lots of ways of applying **Generify!** badly too

[2]( ) Scambler actually shows that Sca interprets ZFC with the Powerset axiom merely removed. However a trivial modification (adding the modal versions of Collection and Separation instead of Replacement) to his system gets you full ZFC^-, so we state this stronger form of the theorem.

(e.g. by just adding single Cohen reals over and over again). But, if we apply **Generify!** and **Reify!** *just right*, we will get ZFC^- + Count.

So Sca can be thought of as providing a stage theory for countabilist versions of the weak iterative conception. This is a pleasing result, but we will raise some questions for the approach in §9.1.

It just remains to ask what our initial starting worlds are. Given Scambler's axioms, we can start with the empty set (we have to shoehorn in the existence of a world with the natural numbers, but this is par for the course). However we could equally think of building up over (for example) a model of ZFC with some classes added on top. Much as with the iterative conception in general, we have a choice as to what to start with. Our two set forming operations are reifying classes and adding in forcing generics, and we can start these processes over lots of different starting domains. Thus we have a plausible stage theory for the countabilist and advocate of the **Forcing Saturation**. And it does so in such a way that ZFC^- + Count is motivated, whilst still explaining why **Universality** fails and **Indefinite Extensibility** holds (there are always new sets at additional stages, and neither reifying the classes of a stage nor adding a forcing generic will allow you to form a set of all sets). So, as with the strong iterative conception, we get **Paradox Diagnosis** here too.

7.4 ( ) Doing without Reify!

It's worth mentioning here that one does not *need* the vertical modality in order to get a conception of stage that will motivate ZFC^- + Count. Although not intended for this purpose (his focus is more linguistic) John Steel has proposed a theory of worlds and sets that will do the job without needing a vertical modality. He proposes (in [Steel, 2014]) a two-sorted theory with variables for sets $x_0, x_1, \ldots$ and variables for universes $W_0, W_1, \ldots$ with the following axioms (here I follow the presentation in [Maddy and Meadows, 2020]):

Definition 52. *Steel's Multiverse Axioms* are as follows:

(i) The axiom scheme stating that if W is a world, and ϕ is an axiom of ZFC, then ϕ holds at W.

(ii) Every world is a transitive proper class.

(iii) If W is a world and $\mathbb{P}$ is a forcing partial order in W, then there is a universe W' containing a generic for W.

(iv) If U is a world, and U can be obtained by forcing over some world W, then W is also a world.

(v) If U and W are worlds then there are G and H that are generic over them such that $U[G] = W[H]$.

A discussion of these axioms, explicitly making the link with countabilism, is available in [Meadows, 2015]. But note we can also think of these as providing a stage theory of a sort for countabilist versions of the weak iterative conception. We start with some proper class model(s) of ZFC, and the set forming operation is *just* **Generify!**.

Formally, we can provide the following axioms:

Definition 53. SteMMe (for **Ste**el-**M**addy-**Me**adows) comprises the following axioms in $\mathscr{L}^{\Diamond}_{\prec,\in}$

(i) Classical first-order logic.

(ii) Extensional plural logic.

(iii) The axiom 'The ordinals do not form a set'.

(iv) Classical S4.2 with the Converse Barcan Formula for every modality.

(v) The necessity of distinctness and stability axioms for $\in$ and $\prec$.

(vi) First-order ZFC.

(vii) The potentialist translations of Separation and Collection.

(viii) **Possible Set-Generics.** The axiom 'If $\mathbb{P}$ is a forcing partial order and $\mathcal{D}$ is a set of dense sets of $\mathbb{P}$, then it's possible that there is a filter meeting each dense set that is a member of D'.

SteMMe suppresses some details of Steel's proposal (in particular, I've ignored moving to ground models of forcing extensions for simplicity). The thought behind it is that at an initial stage we're given some proper class model(s) of ZFC. There is no **Reify!** operation, the only set forming operation is **Generify!**. Still, we can note:

Fact 54. SteMMe interprets ZFC^- under the potentialist translation.[3]

[3] Here's a sketch of the proof:

Proof. (Sketch) Since every world satisfies ZFC many of the axioms are trivial. Moreover SteMMe includes the potentialist translations of Collection and Separation. And clearly SteMMe proves the potentialist translation of Count by **Possible Set-Generics**. □

I conjecture that including the potentialist translations of Collection and Separation in SteMMe is redundant. The consistency proof in [Steel, 2014] for Steel's multiverse axioms also works for SteMMe. There, Steel finds a model for his multiverse axioms by taking a model of ZFC, and adding a generic G for the $Col(< Ord, \omega)$ extension, and letting worlds be of the form $V[G \restriction \alpha]$ (i.e. when we restrict G to α). Letting accessibility be given by forcing, it's easy to see then that this forms a Kripke model for SteMMe (and, in particular it validates S4.3).

Note that (in stark contrast to the strong iterative conception) worlds are proper classes. There is a possible puzzle here—why can't we collect together the sets from one of these proper class worlds to form a set? After all, all the members of some proper classes (e.g. the ordinals) are 'available' for collection at every world. The answer is that the collection forming operation—set forcing—does not allow them to be collected. So we still have a **Paradox Diagnosis** (though one that merits some serious philosophical scrutiny). Although there are worlds containing proper classes, we avoid contradiction by having a suitably 'weak' operation of set formation.

There are many details to be ironed out with these proposals (I will discuss some in Chapter 9). For now it suffices to note that though they are somewhat nascent, there are theories like Sca and SteMMe that provide a modal stage theory for the weak iterative conception that validates **Forcing Saturation** and Count. We provide some assessment of these approaches in Chapter 9. But we might ask at this point, which is better out of the strong iterative conception and the forcing-saturated weak iterative conception?

Chapter 8

Mathematics under the different conceptions

This chapter will examine whether one of the strong iterative conception or the above countabilist versions of the weak iterative conception is best. We'll do this by looking at how mathematics is interpreted under each conception of set, and examine each with respect the theoretical virtues we discussed in Chapter 2. We'll first provide an explanation of how each handles mathematics, before contrasting them side-by-side with respect to our theoretical virtues.

8.1 Mathematics and the strong iterative conception

Let's first recap the situation with the strong iterative conception. As we noted in Chapter 2, ZFC and the strong iterative conception does an extremely good job of interpreting mathematics. A couple of extra things should be mentioned though at this point.

One core problem for the advocate of the strong iterative conception is to resolve questions about **Theory of Infinity**. For, whilst they do have that ZFC is true, ZFC tells us vanishingly little about the behaviour of infinite sets, and in particular the values of the continuum function $f(\aleph_{\alpha+1}) = 2^{\aleph_\alpha}$ or whether large cardinal axioms hold. More has to be done to substantiate new axioms for set theory, and there's a rich literature on the topic.[1]

One kind of mathematics that the advocate of the strong iterative conception has to interpret are the countabilist stage theories. On her view, these stage theories can be concerned with the hereditarily *countable* sets (i.e. sets built up only from countable sets—formally we say that a set is hereditarily countable if

[1] See, for example [Maddy, 1988a], [Maddy, 1988b], [Koellner, 2014], and [Incurvati, 2017] among many others.

it is a countable set containing only hereditarily countable sets). In this way, the advocate of the strong iterative conception holds that the theorist advocating the forcing-saturated weak iterative conception can be interpreted as talking about structures that miss out a great many *large sets* (and in particular all the uncountable ones).

8.2 Mathematics under forcing-saturated weak iterative conceptions

Things are a little more challenging under the forcing-saturated weak iterative conceptions. Because we don't have **Powerset**, we can't just piggy-back off the 'standard' account of mathematics available under the strong iterative conception.

We've seen two versions of the weak iterative conception (given by SteMMe and Sca) that validate **Forcing Saturation**. However, in this context we don't have the Powerset Axiom, and hence can't build many of the usual representations of structures that we want. So there's a number of questions we can ask about the forcing-saturated countabilist interpretation of mathematics:

(1.) How should we understand the study of theories based on ZFC?

(2.) What does 'mainstream' mathematics look like under this conception?

(3.) What does our **Theory of Infinity** look like?

How should we understand the study of theories based on ZFC?

What becomes of our study of ZFC on this approach? The quick answer is that you can still have ZFC *in a model* you just can't have *all subsets* of the sets in those models (since for any set x, there's a collapsing function from x to ω). If you want to have 'uncountable sets' you just have to leave out the subsets that witness bijections with the natural numbers.

(A parenthetical remark that should be included at this point: The idea that sets might be small but 'appear' large in some model appears in the work of Skolem, especially [Skolem, 1922]. Often, however, Skolem's position is cashed out via a scepticism and/or referential indeterminacy by asking the question "How do I know I'm not living in/speaking about a countable model?". The present family of views does not have this flavour, and can instead that we can perfectly well refer to the universe, it is just that the level of **Forcing Saturation** is so strong that we can only talk about 'uncountable' sets by missing out functions.)

One can have very natural looking models here. For example, as well as countable transitive models, it's possible to have transitive models of ZFC containing all ordinals (so called 'inner models') within a model of ZFC^- + Count. Recall that, for example, Sca interprets ZFC under the vertical modality. So any countabilist theory based on Sca will have inner models of ZFC.

Aside from the stage theory, there are also natural axioms that get us inner models for ZFC plus large cardinals. Since the axioms are somewhat complex, I'll provide them in a -box:

( ) I'll mention some in passing, but I won't go into details since the mathematics starts to get tricky. The interested reader is directed to [Barton and Friedman, MS] for further references and a fuller discussion of these examples. One way is to assert the existence of 'sharps'—these imply that there are self-embeddings from many inner models and can be used to get ZFC plus large cardinals in inner models within ZFC^- + Count.[a] Another (related) kind are *axioms of definable determinacy*. Many of these statements (e.g. Projective Determinacy) can be (schematically) rendered in ZFC^- + Count, and also imply that there are inner models of ZFC plus many large cardinals. Finally in [Barton and Friedman, MS] we propose an axiom (the *Ordinal Inner Model Hypothesis*), which implies that every set is countable but also that ZFC with large cardinals added holds in inner models (for the cognoscenti—one can get $0^\sharp$).

[a] See Regula Krapf's PhD thesis [Krapf, 2017] for details of handling sharps in the countabilist context.

There is thus a kind of 'symmetry' between the strong iterative conception and the forcing saturated weak iterative conception. Under the forcing-saturated weak iterative conception, the theories motivated by the strong iterative conception should be understood as holding in transitive models that *miss out subsets* (in particular all the collapsing functions). But under the strong iterative conception, the theories motivated by the forcing saturated weak iterative conception seem to *miss out large sets* (in particular all the uncountable ones).[2]

Mathematics for the countabilist

The picture of mainstream mathematics is much different when we have **Forcing Saturation**. Whilst arithmetic remains unchanged (one can have V_ω exactly as under the strong iterative conception), there are no uncountable set-sized structures. Rather, *all uncountable collections are proper-class-sized*. The study

[2] In [Barton, MS] I've argued that this symmetry can be used to claim that *un*countabilism is in fact *restrictive*.

of *all the real numbers* thus becomes the study of a large proper class.[3] Since there are exactly continuum-many continuous functions between the reals, we can also think of the study of all continuous functions $f : \mathbb{R} \to \mathbb{R}$ as examining a proper class. But whilst the real numbers and class of all continuous functions are proper classes, yet higher mathematics for larger uncountable cardinals cannot be interpreted as about the sets without the use of even higher-order logic. For example the classical study of the space of *all* functions $f : \mathbb{R} \to \mathbb{R}$ (a key structure for functional analysis), cannot be interpreted even by a proper class. One might ask oneself at this point, whether this is *bad* or just *merely different*. We'll return to this issue below (§8.3).

What does our Theory of Infinity look like?

How is **Theory of Infinity** handled? There are (at least) two different kinds of question one could ask:

(1.) How should we understand the **Theory of Infinity** provided by ZFC?

(2.) What is the **Theory of Infinity** simpliciter?

The former question is easily handled under the forcing-saturated weak iterative conception. Since ZFC is only true *relative to a model* that misses out sets, the behaviour of the continuum function (as well as other independent sentences) should be understood via the diverse world-to-world information we get out of the different models of ZFC. This has affinities with some so called 'multiverse' views in the philosophy of set theory (we'll discuss these later in Chapter 9, for our purposes now one can simply read 'multiverse' as the collection of all countabilist stages). For example, Joel-David Hamkins writes:

> ...the continuum hypothesis is a settled question; it is incorrect to describe the CH as an open problem. The answer to CH consists of the expansive, detailed knowledge set theorists have gained about the extent to which it holds and fails in the multiverse, about how to achieve it or its negation in combination with other diverse set-theoretic properties. [Hamkins, 2012, p. 429]

Since there is no maximal ZFC structure for the forcing-saturated weak iterative conception, we have an answer to the question of CH behaves in ZFC set theory. Simply put, it is to be found in how CH behaves across structures that satisfy ZFC. No further answer is needed or possible.

[3] () In fact, since you can think of a real number as coding a countable set, the study of set theory is in a way just the study of real numbers under ZFC^- + Count. This is supported by the fact that second-order arithmetic and ZFC^- + Count are bi-interpretable. See §5.1 of Regula Krapf's PhD thesis [Krapf, 2017] for a nice presentation of this result.

This answer only concerns the impoverished ZFC models for the countabilist. So what is their **Theory of Infinity** simpliciter? This question is answered for *sets*—every set is either finite or countably infinite. So, in a sense, the countabilist has a comprehensive (albeit slightly boring) answer for the relative sizes of *sets*. However, there are still some interesting questions to be had. Since the continuum is a proper class, CH is now a claim about what *proper classes* exist coding bijections between *classes* of sets and the universe. Is every *class* of reals either countable or the size of the universe? This is the open question that the countabilist must address.

( ) One very interesting fact is that in this context CH is *equivalent* to the claim that the universe is bijectable with the ordinals. So we have an immediate link with CH and versions of Global Choice. Moreover, CH is equivalent for the countabilist to the 'limitation of size' principle that all proper classes are the same size.[a] *If* the advocate of the forcing-saturated weak iterative conception could motivate the principle that all proper classes are the same size, they would then have a complete story about **Theory of Infinity**, every set is either (a) finite, (b) infinite, or (c) proper-class-sized, and the continuum hypothesis (rendered as a claim about proper classes) is true.

[a] See here [Holmes et al., 2012], §3.4.

8.3 Contrasting the two conceptions

Is one of the two conceptions better? Both have different ways of responding to **Theory of Infinity** and advocate very different responses. Both have some open questions to answer.

This all raises a question of what will become of the different conceptions, especially when we bear in mind the criteria outlined in Chapter 2. I won't come down one way or the other here—I think there are many questions to be left open for the future. The main point I won't to press is the following: Both are *attractive* conceptions of set.

I do think it's pretty clear that the strong iterative conception, with the rich understanding we have of it and theories motivated on its basis, is well in the lead in the race. This is to be expected, we've only recently starting looking seriously at the forcing-saturated weak iterative conception, and so the strong iterative conception had an enormous head start (a good 50 years or so). Races that seemed one-sided can get more competitive over time though. For example the logical conception is experiencing something of a resurgence due to its possible application in formal semantics having previously been regarded as almost dead-in-the-water (or at least deeply problematic).[4] So it's worth thinking

[4] See [Linnebo, 2006], [Linnebo and Shapiro, F], and [Roberts, MSa].

of how each responds to the desiderata outlined in Chapter 2, contrasting the two, and considering whether the forcing-saturated weak iterative conception might catch up. For the sake of ease, we repeat our theoretical virtues here:

Generous Arena. Find *representatives* for our usual mathematical structures (e.g. the natural numbers, the real numbers) using our theory of sets.

Shared Standard. Provide a standard of correctness for proof in mathematics.

Limits of Thought. Set theory provides a natural place to examine where the limits of human thought are, pushing the boundaries of what might be realistically expected to be known, and exploring where they may finally give out.

Testing Ground for Paradox. Set theory is very *paradox* prone, both in terms of the principles that can be formulated within set theory and when combined with certain philosophical ideas (e.g. absolute generality and mereology). In this way, set theory provides a *testing ground* for seeing when and how ideas explode.

Metamathematical Corral. Provide a theory in which metamathematical investigations of relative provability and consistency strengths can be conducted.

Risk Assessment. Provide a degree of confidence in theories commensurate with their consistency strength.

We also added:

Paradox Diagnosis. Explain why the paradoxical collections aren't sets and which conditions determine sets (and which don't).

Generous arena is handled very differently by the two approaches. But each has their own answer. The strong iterative conception can essentially piggy-back off the standard account of **Generous Arena** given in Chapter 2. Little more needs to be said here.

The case of the forcing saturated weak iterative conception is more controversial. Here the reals are a *proper class* (at least in the non-modal theory). Set theory here is directly akin to second-order arithmetic, and analysis can be thereby interpreted (so long as we allow talk of proper classes). But third-order arithmetic is out of reach, standardly interpreted. However, since we have ZFC plus large cardinals in inner models, proofs using resources from third-order arithmetic and above can be interpreted in *restricted contexts*. Whether this constitutes a hobbling of mathematical practice or just a different approach is a question I leave open for philosophical examination.[5]

This has implications for **Shared Standard**. Both the strong iterative conception and forcing-saturated weak iterative conception provide their own

[5] See [Barton and Friedman, MS] for some more detail.

Generous Arena, and hence their own account of when a proof is legitimate. Each standard is very different though, if we have **Forcing Saturation**, third-order resources are not legitimate for reasoning about the reals. So both have an account of **Shared Standard**, but the forcing-saturated weak iterative conception deviates substantially from the currently accepted norm. This said, under this countabilist approach, proofs in third-order arithmetic and/or ZFC are not *wrong*, they just need to be interpreted in *restricted contexts*. Again, I leave it open whether or not this should count *against* the position or it is simply merely *different*.

Regarding the **Limits of Thought**, both are able to handle Gödelian incompleteness in much the same way (claims about relative provability can be construed as claims about first-order arithmetic, and the first-order arithmetic provided by the two conceptions are not significantly different[6]). However since both provide very different pictures of the role of the continuum and independence, they provide quite different answers to the question of our knowledge of the continuum. The strong iterative conception has several questions to answer about large cardinal independence and the the behaviour of the continuum function. The forcing-saturated version of the weak iterative conception, on the other hand, answers basically all questions about *sets*. Every set is countable, and there are *no* large (or even uncountable) cardinals, even if there are large cardinals and uncountable cardinals in inner models. However, the continuum hypothesis is pushed to a question about class theory, and in particular is connected with global well-orders for the universe (whether there's a proper-class-sized bijection $F : V \rightarrowtail\!\!\!\!\rightarrow Ord$) . As we noted above, if such a countabilist can motivate the claim that all proper classes are the same size, then CH is solved too. But perhaps one can argue that whilst the sets are relatively easily known, the continuum/proper classes are not, and so we leave this question open. But there are at least *avenues* for making philosophical progress on this question.

Moreover, both provide interesting perspectives as a **Testing Ground for Paradox**. This is in two ways. First, the incompatibility between **Powerset** and **Forcing Saturation** and the two conceptions we've discussed provides for an interesting kind of 'paradox' in its own right (this is part of what was at play in the Cohen-Scott Paradox). Interestingly, although each denies the full generality of the other's principles, one can incorporate *partial amounts* thereof. The proponent of **Powerset** can add in limited amounts of **Forcing Saturation**, for restricted kinds of partial order and families of dense sets (this yields a class of axioms known as *forcing axioms*). Interestingly, the addition of such restricted **Forcing Saturation** into the strong iterative conception tends to yield a reso-

[6]Really, all one gets is that the different theories proposed will yield more/less information about the natural numbers. But any theory of arithmetic compatible with one conception is compatible with the other.

lution of CH in the negative, with $2^{\aleph_0} = \aleph_2$.[7] It is not known how to generalise these axioms for higher values of the continuum function. For the proponent of the forcing saturated weak iterative conception of set, we can begin by noting that the existence of uncountable cardinals are a bit like large cardinals—they assert the existence of sets closed under various kinds of operation. For example, the least uncountable cardinal can be thought of as a set that is closed under the formation of hereditarily countable cardinals. Over ZFC$^-$, an uncountable cardinal behaves a bit like an inaccessible cardinal does in ZFC.

> ( ) For example, let κ be the least inaccessible and ω_1 be the least uncountable cardinal. Both are regular, and both provide a natural model for the base theory—V_κ provides a model for ZFC (in fact *second-order* ZFC), and $H(\omega_1)$ provides a model for ZFC$^-$.

Moreover, one *can* postulate the existence of sets with closure under countabilism (just not enough to get you an uncountable cardinal). Here's a slightly tricky example:

> ( ) Consider the following schematic reflection principle (for any ϕ in the language of set theory):
>
> $$\forall x \exists a (x \in a \wedge \text{`}a \text{ is transitive'} \wedge \phi \leftrightarrow \phi^a)$$
>
> i.e. for any set x there is a transitive set a such that $x \in a$ and ϕ is absolute between a and the universe. ZFC$^-$ with this added is known as ZFC$^-_{Ref}$. This theory is very weak—still far below the consistency strength of ZFC (and so is consistent if ZFC is). But it adds in sets with *closure*, in particular if ϕ holds in the universe then ϕ holds restricted to some transitive set a. And since the universe exhibits various closure properties, this version of reflection will imply that there are sets with those closure properties too.

So whilst we know that we'll have to get rid of one of **Forcing Saturation** or **Powerset**, whichever way we go, we can add back in *some restricted versions* of the one we rejected.

Metamathematical Corral is handled immediately. Both conceptions motivate theories that can handle talk of set-theoretic models easily, and so there is no particular difference here. Similarly for **Risk Assessment**, whilst there might be small fluctuations dependent upon which theory is eventually picked, both conceptions can motivate theories with a good deal of strength on an independently plausible conception. We also might think that there's no need to settle on a single conception for **Risk Assessment**, so long as the conceptions

[7] For example the *Proper Forcing Axiom* implies that $2^{\aleph_0} = \aleph_2$. For a survey of the Proper Forcing Axiom, see [Moore, 2010].

seem cogent and coherent, we can have confidence in the consistency of theories that are proved consistent on each picture. In particular, if a theory U is proved consistent by theories motivated under each conception, then more power to U—its consistency is converged upon by two distinct cogent conceptions of set.

For these reasons I think that both the strong iterative conception and the forcing-saturated weak iterative conception are each viable conceptions of set. The strong iterative conception clearly fits better with current orthodoxy, but that's not a good reason to discount the forcing-saturated weak iterative conception out of hand. In the end, I think that a careful analysis is needed, either to choose one of the two or to learn to live with the pluralism they offer. For this to be done successfully, more development of these two (and other) conceptions is required, especially on the side of the juvenile weak iterative conception.

Chapter 9

Conclusions, open questions, and the future

A short summary of what I've argued in this book: I think that set theory provides an interesting case study and tool for both philosophers and mathematicians. I think that by progress in set theory often involves trading off different principles (e.g. **Universality** and **Indefinite Extensibility**, **Powerset** and **Forcing Saturation**). I think that this is the situation we find ourselves in now (at least to some degree).

This said, there's a *lot* more research to be done in this direction. Some areas I have already identified, but some are new and so I want to close with a summary and consolidation of what I take to be the most important questions for moving forward. It will also be helpful to present some objections to what I've argued here, and mention how they could be answered. This will make this 'conclusion' longer than usual, and I hope the reader will indulge me in this.

9.1 The weak iterative conception needs work

Earlier (Chapter 7) I remarked that versions of the strong iterative conception were further ahead in the race as compared to other versions of the weak iterative conception (in particular countabilist ones). Instead, one might argue, they have not qualified to make the start line.

There are a few reasons one could give to substantiate this claim. The strong iterative conception, one might contend, is well-developed. We have an account of what the worlds are (the V_α). By contrast weak iterative conception seems rather underspecified, and clearly in need of sharpening by a further conception. But what are the constraints here? What is to count as a legitimate process? These are all left unanswered by the weak iterative conception and we might worry that the weak iterative conception is not sufficiently well-formulated to provide enough constraints.

Here's a somewhat silly example of a description of an iterative process.

Definition 55. (Informal) The *trivialising conception* of set holds that sets are formed in stages. There are just two stages. At stage 0 we have nothing. At stage 1 we perform the following operation "Form all the sets!". There are no other stages.

What's wrong with this as a version of the weak iterative conception? I think it's important to recall (Chapter 3) what we want out of a conception of set. We want a conception that does the following two things:

(1.) It should motivate a 'good' theory of sets, where "good" is to be spelled out via the theoretical virtues discussed in Chapter 2 (and, indeed, possibly others).

(2.) In particular, it should provide a **Paradox Diagnosis**.

This trivialising conception does not perform well here. In particular it doesn't explain why paradoxical collections don't get into its second stage (since we have no explanation of why the operation **Form all sets!** doesn't form paradoxical ones) and it is totally uninformative about the theory we should adopt. So, yes, it is a legitimate version of the weak iterative conception. But it is also *rubbish*. We can thus safely kick it to the kerb. By contrast the countabilist versions of the weak iterative conception, with their attendant axioms and stage theory, look promising, even if slightly less developed than the strong iterative conception.

That's not to say that there aren't some important questions here that need to be answered under for these countabilist conceptions. An important issue is to work out the details of the countabilist stage theory for the weak iterative conception. One of the major differences between the strong iterative conception and these is that the stage theory of the former is pretty much fully worked out, whereas it is less clear for the latter (though there are options as discussed in Chapter 7). I want to make a few points about moving forward with the project of isolating appropriate stage theories, and the challenges that need to be overcome in order to solidify them as genuine contenders, rather than an up-and-coming prospect.

First, I think that the weak iterative conception is *extremely broad*. This is evidenced by the fact that the trivialising weak iterative conception is a legitimate version of it, even if *terrible* as a conception of set. Moreover, there are very many disparate conceptions that also fall under this banner (e.g. the constructibilist conception and the forcing-saturated conception don't seem to have a whole lot in common beyond their weak iterativity). So I don't think we are going to get a *lot* of informativeness out of the weak iterative conception alone.

However, one thing we *do* get is the idea that there be some sort of description of the universe as unfolding as part of a modal process. And I think the following *is* true: *Legitimate processes should be well-founded.*

Here lies the a challenge for coming up with a more detailed account of the stages for the weak iterative conception: Many of the possible candidates for modal stage theories considered in Chapter 7 are *not* well-founded in the sense that they don't have a well-founded accessibility relation. The problem concerns forcing: It's pretty rare to have forcing models that are the minimal kind of extension under inclusion.

> ( ) For example, imagine I add a single Cohen real G to a stage S to form the stage $S[G]$. By doing so, I can immediately see an infinite descending sequence in the accessibility relation (indeed, I can see a dense ordering of such possible stages), because whenever I add a Cohen real G, there are infinitely many other Cohen reals definable from G, some of which can't be used to define G.

Is this knock-down? I think not. The point is that although *accessibility* is non-well-founded, the notion of a *process* is not.

For a simpler example, suppose I'm given a line segment s in Euclidean space. Now consider the 'modal stage theory' of what I could get by extending s in a single direction. This is an idea of a kind of possible processes that I could do, for any length m I could extend s by m (denote this by "$s^\frown m$"). But any time I do so, I can then see a whole bunch of worlds, indeed a dense sequence, between s and $s^\frown m$ (e.g. $s^\frown \frac{m}{2}$, $s^\frown \frac{m}{3}$, ..., $s^\frown \frac{m}{n}$, ...) that I could have extended to instead. But this doesn't threaten the legitimacy of the procedure, I performed *one* action—*extending* s *by* m. This has been recognised since at least the time of Euclid and Aristotle (indeed, there is a more-than-superficial resemblance to Zeno's dichotomy paradox). What we have in such a case is that the modal *accessibility relation* does not exactly match the kinds of *procedure* we can do. What I suggest is that one looks at the well-founded *subrelations* of the accessibility relation. These will be legitimate possible iterations of 'processes' for both the modal line extension case and a stage theory that incorporates forcing as a method of extension. For the strong iterative conception, it is just their luck that their accessibility relation is well-founded and *matches* their specification of the processes involved in their version of the weak iterative conception. But this needn't be the case. It's then open to us to say that, whilst I can force to a world (and thereby see a descending sequence in the accessibility relation), the way I *get to* any world has to be doable in a well-founded way. But this suggestion, though promising, is very far from being worked out in detail, and represents a substantial open question that needs to be answered for stage theories like Sca and SteMMe. So we ask:

Question. Is there an account (possibly formal) of the weak iterative conception that makes clear the notion of a 'legitimate process'?

In this regard we can also ask:

Question. What other kinds of process are there for underwriting weak iterative conceptions of set?

9.2 Going back the other way

Supposing that this can be worked out, there is a further hole to patch. In the case of the strong iterative conception and ZFC, we have the pleasing result that not only does the modal stage theory motivate ZFC, but ZFC establishes the correctness of the stage theory. In particular we have:

Theorem 30 (ZF) For every set x there is an ordinal α such that $x \in V_\alpha$.

So we can go both ways, the strong iterative conception (suitably formulated) can motivate ZFC, and ZFC (indeed ZF) can recover a notion of stage and prove that every set is a member of a stage.

The countabilist versions of the weak iterative conception that we've discussed are not in quite such a rosy state. Whilst we can motivate ZFC$^-$ + Count, it is unknown whether or not there could be a related theory in which we can prove an analogous theorem. We therefore ask:

Question. Is there a reasonable presentation of a stage theory S that motivates an extension T of ZFC$^-$ + Count, but where one can (in T) recover the stages of S and prove that every x is a member of some stage?

9.3 Potentialism, actualism, and absolute generality

Throughout this book, we've been discussing modal stage theories. An important question in the philosophy of mathematics concerns how we should think of these modalities. I want to put on the table three possible answers to this question:

Actualism/Universism. There is a single universe of sets and a definite plurality of all sets.

Potentialism. There is a single universe of sets, but it is modally indefinite. There is no definite plurality of all sets.

Multiversism. There is no single universe of all sets, rather many universes.

(**Note:** It may be that we should relativise these questions to a given conception, with different questions of how the stages are interpreted for different conceptions.)

These views do not exhaust the logical space (e.g. we could have a universe that is indefinite, but not modally so, e.g. [Feferman, 2010], [Scambler, 2020]) but they are the main ones that are relevant for stage theories. Each view suggests a different way of philosophically interpreting the relevant modalities. Let's start with the modalities involved in the stage theories we've considered. The actualist regards the use of modality as a mere heuristic for talking about the stage-theoretic structure of the universe. The potentialist takes the modality seriously, and thinks that it is somehow indicative of the fundamental nature of reality. The multiversist also thinks that the modality is a mere heuristic but in a very different way from the actualist, for them it is a way of talking about interrelationships between the different universes on offer, and ways of moving between them.

Each view has it's own idiosyncrasies and suite of problems to be addressed. One aspect of each is how we regard the *determinacy of truth* concerning mathematical claims (in particular in the language of set theory). The universist will likely assert that every sentence of set theory has a definite truth value—assuming we can refer to their universe without issue, the truth or falsity of claims should just be understood as the truth or falsity of claims there. Likewise the multiversist will likely assert that there are set-theoretic claims of indeterminate truth value—true in some worlds and false in others. The potentialist (given mirroring) is likely to fall on the side of determinacy, at least insofar as 'normal' mathematical claims go (which should be understood under the potentialist translation).

For the universist, there is the old problem of the nature of proper classes. For example, Øystein Linnebo writes:

> Since a set is completely characterized by its elements, any plurality...seems to provide a complete and precise characterization of a set... What more could be needed for such a set to exist?[1] [Linnebo, 2010, p. 147]

The problem is as follows. Given the stages of any version of the weak iterative conception, the universist holds that there is a determinate totality of all the sets in the stages. This can be cashed out in plural terms; there are some sets xx such that that there is no set of all the xx (for ease, let's just assume that the xx comprise every pure set). But what is it then that stops us forming these sets

[1] [Linnebo, 2010] is especially concerned with the semantics of *plural* quantification here, and I've suppressed this detail for clarity.

into a new set? We have a definite plurality of them, and so could characterise the relevant membership relation. One response is to say that contradiction would ensue. But this only holds if you *assume* that the xx contain every possible pure set. So, the universist has to come up with a meaningful explanation of proper classes that makes it clear why they're different from sets, and why the seeming ability to talk about such collections isn't an issue.

Similarly, many see the generality and flexibility of forcing as evidence that a given domain of sets can be expanded. Here's Hamkins on the subject:

> A stubborn geometer might insist—like an exotic-travelogue writer who never actually ventures west of seventh avenue—that only Euclidean geometry is real and that all the various non-Euclidean geometries are merely curious simulations within it. Such a position is self-consistent, although stifling, for it appears to miss out on the geometrical insights that can arise from the other modes of reasoning. Similarly, a set theorist with the universe view can insist on an absolute background universe V, regarding all forcing extensions and other models as curious complex simulations within it. (I have personally witnessed the necessary contortions for class forcing.) Such a perspective may be entirely self-consistent, and I am not arguing that the universe view is incoherent, but rather, my point is that if one regards all outer models of the universe as merely simulated inside it via complex formalisms, one may miss out on insights that could arise from the simpler philosophical attitude taking them as fully real. [Hamkins, 2012, p. 426]

So, an open question for the universist is how we should interpret the use of forcing *over the universe* (including how natural these interpretations are).[2]

As noted above, the multiversist faces no such difficulties. However they find themselves in hot water concerning the usual problems of generality relativism. They assert that there is no absolute universe, but then immediately seem to make claims about *all* universes. The immediate question is: "Why can't we just understand this domain as the absolute universe?". Since the literature here is *enormous*, I'll say no more about it, but merely point out that it remains open.[3]

The potentialist does not face these problems. If one believes that one can always **Reify!** and **Generify!** over any definite plurality, and talk about these processes modally, one does not face the same difficulties. Any definite plurality forms a set, and any definite plurality can be forced over.[4] Since the universe is

[2] This is a literature I've contributed to in [Barton, 2021] and [Antos et al., 2021].

[3] For further reading see [Rayo and Uzquiano, 2006], [Florio and Linnebo, 2021] (esp. Chapter 11), and [Studd, 2019].

[4] There is a question of whether the motivations for these different positions are satisfactory, see [Roberts, MSb].

not modally definite, they may contend that there is no definite plurality to be climbed or forced over. This is the response of both [Linnebo, 2010] (for proper classes only) and [Scambler, 2021] (for both). *Given that their modality is legitimate*, a response can be made out along these lines. An important question is thus whether that modality can be given an acceptable gloss, or seems parasitic on other (unavailable) notions.[5] The multiversist and universist can both explain the modality by reducing it to other notions (direct quantification over universes for the former, restricted quantification over the stages for the latter). So there is a real question of whether the potentialist has just exchanged one suite of problems for another, and whether one set is especially worse.

A final question regarding absolute generality concerns the similarity between the reasoning involved in Cantor-Russell and Cohen-Scott. Some authors have argued that the similarity between the two suggests that if one is a **Reify!** potentialist/multiversist, then one should be a **Generify!** potentialist/multiversist too.[6] Really substantiating this thesis would require a more detailed analysis of the similarities between the two pieces of reasoning, and is an open philosophical problem.

Note: This seems like a difficult issue to address, since any such response will have to distinguish both Cantor-Russell and Cohen-Scott from other kinds of 'diagonal' argument where an 'indefinite extensibility' response is not so attractive (e.g. the halting problem, see [Meadows, 2015]). I do not see an easy way to answer this question, in particular because it is not clear to me if there if there is a sharp characterisation of the notion of *diagonal argument* (perhaps instead it is a more 'family resemblance' concept?).[7]

9.4 Connection to conceptual engineering

One salient point to be noted in what I've argued here is that there is a close link to much of the literature on *conceptual engineering*. This field concerns itself with the evaluation, design, and implementation of our concepts.[8] There are affinities between what we've discussed here and this literature. For example, Kevin Scharp has argued that our naive concept of truth is inconsistent, and should be replaced with two concepts (*ascending truth* and *descending truth*) which validate each direction of the Tarski biconditionals separately, but there is no consistent concept that validates both [Scharp, 2013]. There are clear similarities here with the way in which **Universality** and **Indefinite Extensibility** can be traded off, and how **Forcing Saturation** conflicts with **Powerset**. There is a natural project here to view these moves in the light of conceptual engineering

[5] See [Linnebo, 2018], Chs. 3 and 12 for some discussion.

[6] See [Meadows, 2015], [Scambler, 2021], and [Builes and Wilson, 2022] for discussion.

[7] I thank Toby Meadows for some discussion of this point. See also [Simmons, 1990].

[8] See here [Chalmers, 2020] for a survey.

([Incurvati, 2020] explicitly makes this connection for **Indefinite Extensibility** and **Universality**). So we ask:

Question. Should we, and if so how, view the project of trading off features of concepts/conceptions of set as an exercise in conceptual engineering?

9.5 The story is too neat, and ignores much

In this book, I've presented the idea that we can view different attractive conceptions of set as arising out of trading off **Forcing Saturation** and **Powerset**. But I want to emphasise that whilst I do think this is a fundamental tension, there are *many* more options out there, some of which are weakly iterative. what about, for instance, inner model theory and the Ultimate-L programme [Woodin, 2017]? I won't go into detail about this here, but the rough idea is to come up with a version of L that is able to give a good structure theory for V and still incorporate large cardinals ($V = L$ implies that many large cardinals don't exist). What about other proposals for set-theoretic axioms (e.g. Freiling's darts)? Isn't all this a bit narrow?

Yes! It is absolutely too narrow, and space doesn't permit me to go into the full details of every possible direction in set theory. My point here was not to propose **Powerset** and **Forcing Saturation** as *the* two possibilities for set-theoretic development (though I *do* think they might be *especially* attractive to philosophers). My focus was rather to articulate the idea that in certain contexts we can see conceptions as emerging from trading off inconsistent principles, and thereby highlight some similarities between our own predicament and that of our intellectual ancestors. In particular, I made simplifying assumptions there too—there's far more than the conceptions I concentrated on.

There's many twists and turns we could have taken. But really the space of conceptions should be far broader than these pages indicate and are probably not as conceptually neat as they might be. **Regarding breadth:** I've said little, for instance, about some of the conceptions of set considered in [Incurvati, 2020] like the graph conception or those based on non-classical logic such as paraconsistent (e.g. [Priest, 2002], [Jockwich et al., 2022]) or constructivist/intuitionist logics (e.g. [Feferman, 2010], [Bell, 2014], [Scambler, 2020]), or predicativity (e.g. [Feferman and Hellman, 1995], [Linnebo and Shapiro, F]).

The point is just the following: This book isn't meant to be providing a classification for every conception of set. My point is just that by considering (i) the interrelations between different conceptions, and (ii) how we trade off inconsistent principles, we can come to understand better the space of possibilities for articulating the mathematically fertile notion of collection.

9.6 Plato and friends

The next objection comes from the staunch set-theoretic realist/platonist, who thinks that there's just a world of sets 'out there' where every set-theoretic sentence has a definite truth value. Conceptions of set are great and all, but at the end of the day the theories they motivate are either true or false about this universe, and this is the only arbiter of correctness we need. All this talk of theoretical virtues and conceptions of set is a mere red herring.

I don't find this line of argument very persuasive at all. I think the history of set theory, with all its twists and turns, false starts, and possible choice points, indicates that this just isn't a very fruitful way to look at things. To see this, let's grant for the sake of argument that there is such a platonistic realm. What should we think of our talk concerning it? There is a pessimistic probabilistic argument available here: Do we really think, out of all the possible conceptions we might have and all the ways we might have gone and continue to go, that we will really select the 'right' one? I think it entirely possible what we've discussed here is probably a very small snapshot of what is quite a large space. These may well just be a fraction of all the possible conceptions available to humans and gods. What is the probability (given our lack of perceptual interaction with this universe) that we happen to pick the right conception? I would say low.[9]

One could, as a response, say that we *do* have some sort of perception of the universe of sets. I don't have much to say here, beyond the well-worn point that this seems like mysticism to me. Another option is simply a fatalistic pessimism about our chances. But I see a better way out—to regard the interesting questions as ones concerning what we *do* with our conceptions and the theories they motivate, and how they interact with our knowledge as a whole. This strikes me as an area where we can learn and make progress, rather than simply arguing about whose mystical eye sees the farthest.[10]

9.7 Pluralism?

I've argued that we now find ourselves at a fundamental choice point, do we go with **Forcing Saturation**, **Powerset**, or something else entirely? There is, however, a different option: We might end up in a situation in which the various conceptions perform better with respect to certain criteria and/or in different contexts. It's possible that we might be led to a strong kind of pluralism, where claims using the term "set" need to be relativised to a particular kind of conception in order to be assessed for truth. There's a special challenge for analysing

[9] I also make a version of this argument in more detail in [Barton, 2022].

[10] This way of thinking has *some* affinities with Maddy's naturalism [Maddy, 1997], second philosophy [Maddy, 2007], and thin realism [Maddy, 2011]. I differ from her in that I think that an appealing underlying conception is more than a mere "useful heuristic' [Maddy, 2011, p. 136].

mathematical practice here. Normally (at least within ZFC set theory) the 'spectre' of pluralism does not too radically alter the typing of mathematical objects (e.g. within different theories extending ZFC the reals are always a set). However here we do have significantly different types—the continuum might be a proper class under for the countabilist but a tiny accessible set under the strong iterative conception. To me, it seems philosophically open which route we take, or even if we need to pick *one*. So we ask:

Question. What are the prospects for a set-theoretic pluralism arising out of the different conceptions of set discussed here?

9.8 Not the final word

I hope to have convinced the reader that there's a host of interesting philosophical and mathematical questions to be found within contemporary philosophy of set theory. I want to close with a word on the methodology of progress in this field. We can only hope to make serious advances on these issues by thoughtful and meticulous examination of different conceptions. A full study of these problems will thus require a massive effort from historians, philosophers, and sociologists of mathematics, as well as philosophically interested mathematicians, and so there's a real opportunity for collaboration from people working in many fields. Even then though, it's not clear how much control we have over our semantic whims.[11] It may be that significant *set-theoretic activism* is needed in order to get conceptions accepted as legitimate and under consideration. In this way, though mathematics has its own norms and methods of reasoning, the present study suggests a radical *anti-exceptionalism* about mathematics as contiguous with other human endeavours. The future is open and exciting, with a good deal of work to be done in understanding the world(s) of infinite sets.

[11] The idea that we don't have much control is advocated by [Wilson, 2006] and [Cappelen, 2018].

Bibliography

[Antos et al., 2021] Antos, C., Barton, N., and Friedman, S.-D. (2021). Universism and extensions of V. *The Review of Symbolic Logic*, 14(1):112–154.

[Bagaria, 2005] Bagaria, J. (2005). Natural axioms of set theory and the continuum problem. In *Proceedings of the 12th International Congress of Logic, Methodology, and Philosophy of Science*, pages 43–64. King's College London Publications.

[Barton, 2021] Barton, N. (2021). Indeterminateness and 'the' universe of sets: Multiversism, potentialism, and pluralism. In Fitting, M., editor, *Selected Topics From Contemporary Logics*, Landscapes in Logic, pages 105–182. College Publications.

[Barton, 2022] Barton, N. (2022). *Structural Relativity and Informal Rigour*, pages 133–174. Springer International Publishing, Cham.

[Barton, MS] Barton, N. (MS). Is (un)countabilism restrictive? Manuscript under review. Pre-Print: `https://philpapers.org/rec/BARIUR`.

[Barton and Friedman, 2019] Barton, N. and Friedman, S. (2019). Set theory and structures. In Sarikaya, D., Kant, D., and Centrone, S., editors, *Reflections on the Foundations of Mathematics*, pages 223–253. Springer Verlag.

[Barton and Friedman, MS] Barton, N. and Friedman, S. (MS). Countabilism and maximality principles. Manuscript under review. Preprint: `https://philpapers.org/rec/BARCAM-5`.

[Barton et al., 2022] Barton, N., Müller, M., and Prunescu, M. (2022). On representations of intended structures in foundational theories. *Journal of Philosophical Logic*, 51(2):283–296.

[Bell, 2011] Bell, J. (2011). *Set Theory: Boolean-Valued Models and Independence Proofs*. Oxford University Press.

[Bell, 2014] Bell, J. (2014). *Intuitionistic Set Theory*. Studies in logic. College Publications.

[Boolos, 1998] Boolos, G. (1998). Reply to Charles Parsons' 'sets and classes'. In *Logic, Logic, and Logic*, pages 30–36. Harvard University Press.

[Builes and Wilson, 2022] Builes, D. and Wilson, J. M. (2022). In defense of countabilism. *Philosophical Studies*, 179(7):2199–2236.

[Button, 2021a] Button, T. (2021a). Level Theory Part 1: Axiomatizing the bare idea of a cumulative hierarchy of sets. *The Bulletin of Symbolic Logic*, 27(4):436–460.

[Button, 2021b] Button, T. (2021b). Level Theory, Part 2: Axiomatizing the bare idea of a potential hierarchy. *Bulletin of Symbolic Logic*, 27(4):461–484.

[Caicedo et al., 2017] Caicedo, A. E., Cummings, J., Koellner, P., and Larson, P. B., editors (2017). *Foundations of Mathematics: Logic at Harvard Essays in Honor of W. Hugh Woodin's 60th Birthday*, volume 690 of *Contemporary Mathematics*. American Mathematical Society.

[Cain, 1995] Cain, J. (1995). Infinite utility. *Australasian Journal of Philosophy*, 73(3):401–404.

[Cappelen, 2018] Cappelen, H. (2018). *Fixing Language: An Essay on Conceptual Engineering*. Oxford: Oxford University Press.

[Chalmers, 2020] Chalmers, D. J. (2020). What is conceptual engineering and what should it be? *Inquiry*, 0(0):1–18.

[Cohen, 1963] Cohen, P. (1963). The independence of the continuum hypothesis. *Proceedings of the National Academy of Sciences of the United States of America*, 50:1143–1148.

[Cohen, 2002] Cohen, P. (2002). The Discovery of Forcing. *Rocky Mountain Journal of Mathematics*, 32(4):1071 – 1100.

[Drake, 1974] Drake, F. R. (1974). *Set Theory: An Introduction to Large Cardinals*. North Holland Publishing Co.

[Drake and Singh, 1996] Drake, F. R. and Singh, D. (1996). *Intermediate Set Theory*. Wiley.

[Enderton, 1977] Enderton, H. (1977). *Elements of Set Theory*. Academic Press.

[Feferman, 2010] Feferman, S. (2010). *On the strength of some semi-constructive set theories*, pages 109–129. College Publications.

[Feferman and Hellman, 1995] Feferman, S. and Hellman, G. (1995). Predicative foundations of arithmetic. *Journal of Philosophical Logic*, 24(1):1–17.

[Ferreirós, 2007] Ferreirós, J. (2007). *Labyrinth of Thought A History of Set Theory and Its Role in Modern Mathematics (2nd Edition)*. Springer, Birkhäuser.

[Fine, 2005] Fine, K. (2005). Class and membership. *Journal of Philosophy*, 102(11):547–572.

[Florio and Linnebo, 2021] Florio, S. and Linnebo, O. (2021). *The Many and the One: A Philosophical Study of Plural Logic*. Oxford University Press.

[Giaquinto, 2002] Giaquinto, M. (2002). *The Search for Certainty: A Philosophical Account of Foundations of Mathematics*. Oxford University Press.

[Gitman et al., 2016] Gitman, V., Hamkins, J. D., and Johnstone, T. A. (2016). What is the theory ZFC without power set? *Mathematical Logic Quarterly*, 62(4-5):391–406.

[Gödel, 1940] Gödel, K. (1940). *The Consistency of The Continuum Hypothesis*. Princeton University Press.

[Hamkins, 2012] Hamkins, J. D. (2012). The set-theoretic multiverse. *The Review of Symbolic Logic*, 5(3):416–449.

[Hamkins and Lewis, 2000] Hamkins, J. D. and Lewis, A. (2000). Infinite time Turing machines. *Journal of Symbolic Logic*, 65(2):567–604.

[Hamkins and Loewe, 2008] Hamkins, J. D. and Loewe, B. (2008). The modal logic of forcing. *Transactions of the American Mathematical Society*, 360(4):1793–1817.

[Hamkins and Montero, 2000] Hamkins, J. D. and Montero, B. (2000). With infinite utility, more needn't be better. *Australasian Journal of Philosophy*, 78(2):231–240.

[Holmes et al., 2012] Holmes, M., Forster, T., and Libert, T. (2012). Alternative set theories. In *Sets and extensions in the twentieth century*, volume 6 of *Handb. Hist. Log.*, pages 559–632. Elsevier/North-Holland, Amsterdam.

[Incurvati, 2017] Incurvati, L. (2017). Maximality principles in set theory. *Philosophia Mathematica*, 25(2):159–193.

[Incurvati, 2020] Incurvati, L. (2020). *Conceptions of Set and the Foundations of Mathematics*. Cambridge University Press.

[Jockwich et al., 2022] Jockwich, S., Tarafder, S., and Venturi, G. (2022). Ideal objects for set theory. *Journal of Philosophical Logic*, 51(3):583–602.

[Kanamori, 2009] Kanamori, A. (2009). *The Higher Infinite: Large Cardinals in Set Theory from Their Beginnings*. Springer, 2nd edition.

[Koellner, 2009] Koellner, P. (2009). On reflection principles. *Annals of Pure and Applied Logic*, 157:206–219.

[Koellner, 2014] Koellner, P. (2014). Large cardinals and determinacy. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, spring 2014 edition.

[Krapf, 2017] Krapf, R. (2017). *Class forcing and second-order arithmetic*. PhD thesis, Rheinische Friedrich-Wilhelms-Universität Bonn.

[Kunen, 1980] Kunen, K. (1980). *Set Theory: An Introduction to Independence Proofs*. Elsevier.

[Kunen, 2013] Kunen, K. (2013). *Set Theory*. College Publications.

[Linnebo, 2006] Linnebo, Ø. (2006). Sets, properties, and unrestricted quantification. In *Absolute Generality*. Oxford University Press.

[Linnebo, 2010] Linnebo, Ø. (2010). Pluralities and sets. *Journal of Philosophy*, 107(3):144–164.

[Linnebo, 2013] Linnebo, Ø. (2013). The potential hierarchy of sets. *The Review of Symbolic Logic*, 6(2):205–228.

[Linnebo, 2014] Linnebo, Ø. (2014). Plural quantification. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Fall 2014 edition.

[Linnebo, 2018] Linnebo, O. (2018). *Thin Objects: An Abstractionist Account*. Oxford University Press.

[Linnebo and Shapiro, F] Linnebo, O. and Shapiro, S. (F). Predicativism as a form of potentialism. *Review of Symbolic Logic*, pages 1–32. Forthcoming.

[Mac Lane, 1986] Mac Lane, S. (1986). *Mathematics: Form and Function*. Springer-Verlag.

[Maddy, 1988a] Maddy, P. (1988a). Believing the axioms. I. *The Journal of Symbolic Logic*, 53(2):481–511.

[Maddy, 1988b] Maddy, P. (1988b). Believing the axioms II. *The Journal of Symbolic Logic*, 53(3):736–764.

[Maddy, 1997] Maddy, P. (1997). *Naturalism in Mathematics*. Oxford University Press.

[Maddy, 2007] Maddy, P. (2007). *Second Philosophy*. Oxford University Press.

[Maddy, 2011] Maddy, P. (2011). *Defending the Axioms*. Oxford University Press.

[Maddy, 2017] Maddy, P. (2017). Set-theoretic foundations. In *[Caicedo et al., 2017]*, pages 289–322. American Mathematical Society.

[Maddy, 2019] Maddy, P. (2019). *What Do We Want a Foundation to Do?*, pages 293–311. Springer International Publishing, Cham.

[Maddy and Meadows, 2020] Maddy, P. and Meadows, T. (2020). A reconstruction of Steel's multiverse project. *The Bulletin of Symbolic Logic*, 26(2):118–169.

[Meadows, 2015] Meadows, T. (2015). Naive infinitism. *Notre Dame Journal of Formal Logic*, 56(1):191–212.

[Menzel, 1986] Menzel, C. (1986). On the iterative explanation of the paradoxes. *Philosophical Studies*, 49:37–61.

[Menzel, 2014] Menzel, C. (2014). $ZFCU$, wide sets, and the iterative conception. *Journal of Philosophy*, 111(2):57–83.

[Moore, 1990] Moore, A. W. (1990). *The Infinite*. Routledge.

[Moore, 2010] Moore, J. T. (2010). The proper forcing axiom. In *Proceedings of the International Congress of Mathematicians 2010 (ICM 2010)*, pages 3–29.

[Muller, 2001] Muller, F. A. (2001). Sets, classes and categories. *British Journal for the Philosophy of Science*, 52:539–573.

[Oliver and Smiley, 2013] Oliver, A. and Smiley, T. (2013). *Plural Logic*. Oxford University Press.

[Parsons, 1983] Parsons, C. (1983). Sets and modality. In *Mathematics in Philosophy: Selected Essays*. Cornell University Press.

[Posy, 2020] Posy, C. J. (2020). *Mathematical Intuitionism*. Elements in the Philosophy of Mathematics. Cambridge University Press.

[Potter, 2004] Potter, M. (2004). *Set Theory and its Philosophy: A Critical Introduction*. Oxford University Press.

[Priest, 2002] Priest, G. (2002). *Beyond the Limits of Thought*. Clarendon Press.

[Quine, 1937] Quine, W. (1937). New foundations for mathematical logic. *The American Mathematical Monthly*, 44:70–80.

[Rayo and Uzquiano, 2006] Rayo, A. and Uzquiano, G., editors (2006). *Absolute Generality*. Clarendon Press.

[Reinhardt, 1974] Reinhardt, W. (1974). Remarks on reflection principles, large cardinals, and elementary embeddings. *Proceedings of Symposia in Pure Mathematics*, 13:189–205.

[Rittberg, 2020] Rittberg, C. J. (2020). *Mathematical Practices Can Be Metaphysically Laden*, pages 1–26. Springer International Publishing, Cham.

[Roberts, 2019] Roberts, S. (2019). Modal structuralism and reflection. *The Review of Symbolic Logic*, 12(4):823–860.

[Roberts, MSa] Roberts, S. (MSa). The iterative conception of properties. Unpublished manuscript.

[Roberts, MSb] Roberts, S. (MSb). Ultimate v. Manuscript. Preprint: `https://samrroberts.files.wordpress.com/2022/04/ultimate_v.pdf`.

[Scambler, 2020] Scambler, C. (2020). An indeterminate universe of sets. *Synthese*, 197(2):545–573.

[Scambler, 2021] Scambler, C. (2021). Can all things be counted? *Journal of Philosophical Logic*.

[Scharp, 2013] Scharp, K. (2013). *Replacing Truth*. Oxford University Press.

[Scott, 1977] Scott, D. (1977). Foreword to *boolean-valued models and independence proofs*. In *[Bell, 2011]*, pages xiii–xviii. Oxford University Press.

[Shapiro and Wright, 2006] Shapiro, S. and Wright, C. (2006). All things indefinitely extensible. In Rayo, A. and Uzquiano, G., editors, *Absolute Generality*. Oxford University Press.

[Simmons, 1990] Simmons, K. (1990). The diagonal argument and the liar. *Journal of Philosophical Logic*, 19(3):277–303.

[Skolem, 1922] Skolem, T. (1922). Some remarks on axiomitized set theory. In *A Source Book in Mathematical Logic, 1879-1931, van Heijenoort 1967*, pages 290–301. Harvard University Press.

[Steel, 2014] Steel, J. (2014). Gödel's program. In Kennedy, J., editor, *Interpreting Gödel*. Cambridge University Press.

[Studd, 2013] Studd, J. P. (2013). The iterative conception of set: A (bi-) modal axiomatisation. *Journal of Philosophical Logic*, 42(5):697–725.

[Studd, 2019] Studd, J. P. (2019). *Everything, More or Less: A Defence of Generality Relativism*. Oxford University Press.

[Uzquiano, 2006] Uzquiano, G. (2006). Unrestricted unrestricted quantification: The cardinal problem of absolute generality. In *Absolute Generality*. Oxford University Press.

[Weaver, 2014] Weaver, N. (2014). *Forcing for Mathematicians*. World Scientific.

[Whittle, 2015] Whittle, B. (2015). On infinite size. *Oxford Studies in Metaphysics*, 9:3–19.

[Whittle, 2018] Whittle, B. (2018). Size and function. *Erkenntnis*, 83(4):853–873.

[Wilson, 2006] Wilson, M. (2006). *Wandering Significance: An Essay on Conceptual Behaviour*. Oxford University Press.

[Woodin, 2017] Woodin, W. H. (2017). In search of Ultimate-L: The 19th Midrasha Mathematicae lectures. *The Bulletin of Symbolic Logic*, 23(1):1–109.

[Zarach, 1996] Zarach, A. M. (1996). *Replacement $\nrightarrow$ collection*, volume 6 of *Lecture Notes in Logic*, pages 307–322. Springer-Verlag, Berlin.